

# Praise, Blame, Obligation, and DWE: Toward a Framework for Classical Supererogation and Kin<sup>1</sup>

Paul McNamara  
Department of Philosophy  
University of New Hampshire  
Durham, NH 03824-3574 USA  
paulm@unh.edu

**Abstract.** Continuing prior work by the author, a simple classical system for personal obligation is integrated with a fairly rich system for aretaic (agent-evaluative) appraisal. I then explore various relationships between definable aretaic statuses such as praiseworthiness and blameworthiness and deontic statuses such as obligatoriness and impermissibility. I focus on partitions of the normative statuses generated ("normative positions" but without explicit representation of agency). In addition to being able to model and explore fundamental questions in ethical theory about the connection between blame, praise, permissibility and obligation, this allows me to carefully represent schemes for supererogation and kin. These controversial concepts have provided challenges to both ethical theory and deontic logic, and are among deontic logic's test cases.

**Keywords:** Supererogation, Praise, Blame, Obligation, Deontic, DWE.

## Introduction

I have delineated a framework called DWE ("Doing Well Enough") in [7-11] for modeling the logical structure of fundamental but neglected features of common sense morality. I have focused on (personal) *deontic notions*: notions used to evaluate the status of *acts (exercises of agency)* or *states of affairs vis a vis* how (not just if) they would or would not satisfy a person's obligations. The particular focus in DWE is on a set of notions in the logical neighborhood of that of *exceeding the moral minimum* (cf. "action beyond the call of duty"). However, DWE, like virtually all systems of deontic logic, contains no resources for representing *aretaic notions*: notions such as praiseworthiness and blameworthiness that are used primarily to evaluate *agents*, especially (but not exclusively) for the way in which their agency reflects their worth as persons.<sup>2</sup> Thus there is a gap between this work and that of the traditional work on supererogation and kin in ethical theory from the mid-twentieth century forward. For with no representation of praiseworthiness or blameworthiness, there is no way to represent the standard conception of supererogation, much less that of "offense" (or "supererogation") the purported mirror image of supererogation. This neglect is particularly pressing given DWE, since it purports to model doing more than the minimum or going beyond the call, and if that concept is the same as supererogation, then DWE must fail in its attempt, since it lacks any aretaic elements.

A first step to rectify my neglect was taken in [12], and I will integrate some of that material on aretaic notions with work on personal obligation from [13]. Having an integrated representation of both aretaic and deontic concepts will allow for the representation of a diversity of positions in the ethical theory literature on connections often asserted or assumed to hold between these conceptual domains. Derivatively, the resulting framework allows for the representation of the main prior approaches to supererogation and offense other than my own, for example, one of Chisholm's approaches, the earlier Meinong-Schwartz approach, what Mellema calls the standard account, and Mellema's own extension of the standard account. I will discuss the latter three

---

<sup>1</sup> This is an expanded version of a paper presented at DEON'08, and published as "Praise, Blame, Obligation, and Beyond: Toward a Comprehensive Framework for the Classical Conception of Supererogation and Kin", in *Deontic Logic in Computer Science*, Ron van der Meyden and Leendart van der Torre (eds), Lecture Notes in Artificial Intelligence, 5076 (Berlin: Springer Verlag, 2008).

<sup>2</sup> "Aretaic" in ethical theory stems from the Greek term, *arête*, for human excellence or virtue. However, it is used here for moral evaluation of agents generally, and not just on the basis of their (non-transient) character traits.

here. The logical framework also sheds light on various issues in the traditional literature on supererogation. In the end I will show that the classical analysis of supererogation is fundamentally flawed in a way that is both illuminating and ironic.

In section 1, I give a quick sketch of the DWE framework and then explain why developing a logic for aretaic appraisal is particularly pressing in light of DWE. In section 2, I introduce a simple modal logic for what is predetermined (necessary) for an agent, and then in Andersonian-Kangerian style, generate a simple classical deontic fragment that includes Standard Deontic Logic (SDL). In section 3, I introduce FAA, a "framework for aretaic appraisal", a modification of that sketched in [12], which allows for the definition of a variety of aretaic notions, and I characterize two strengthenings of this framework,  $FAA^C$  and  $FAA^{CNC}$ . In section 4, I characterize some partitions  $FAA^{CNC}$  generates, leading to a 7-fold partition of aretaic normative positions.<sup>3</sup> In section 5, I combine this aretaic partition with the classical deontic partition from section 2, which generates a potentially 21-fold partition of combined deontic-aretaic normative positions. I explore two *prima facie* plausible bridging principles, that what is praiseworthy for an agent is always permissible and that what is blameworthy for her is never obligatory, and identify the eliminations of normative positions these reductive principles would entail. I also raise substantive doubts about these at-first-glance very plausible principles, and suggest that agent appraisal and act appraisal often come apart enough that the apt aretaic appraisal can even have a non-neutral value (positive or negative) and the apt deontic appraisal can have a contrary non-neutral value. In section 6, I turn to what I call the classical analysis of supererogation and offense, and to Mellema's addition of the notions of quasi-supererogation and quasi-offense, and I identify the six places where these fall on the prior aretaic-deontic partitions. I also briefly consider the Meinong-Schwartz aretaic-deontic ranking thesis, as well as Meinong's "laws of omission". I then examine some additional stronger simple aretaic-deontic bridging principles, that whatever is praiseworthy is obligatory, and that whatever is blameworthy is impermissible, as well as their converses, and show how these lead to greater reductions in our 21-fold partition, especially with an eye to how they impact the classical analyses of supererogation and kin. I go on to suggest that one simple particular thesis might be behind the widespread skepticism about offenses even among friends of supererogation, but I also argue that the thesis is dubious, however plausible it is at first blush. Indeed, I think the reflections challenge one long standing line of argument for rejecting offenses, to the benefit of ethical theory. In section 7, I show that the traditional analysis of supererogation is fundamentally flawed, and that the resources of DWE help to reveal this previously overlooked fact. I then argue that the classical conception of supererogation presupposes the concept from DWE of *doing more than one has to do* (but not vice versa), making DWE more basic.

In all this, my intention is two-fold: to formally model neglected normative statuses, especially agent-evaluative ones here, and correlatively, to counter the deeply entrenched objection (bias?) from ethical theorists that deontic logic is utterly irrelevant to their enterprise and is a dismal failure in that regard. Indeed, I would submit that some of the work herein is *ahead* of the ethical theory curve in that regard, and that frameworks like the simple one below place questions in sharp relief that ought to clearly benefit ethical theorists.

## 1 Doing Well Enough (DWE)

### 1.1 Syntax

Suppose that I am obligated to be the first to convey some slightly delicate information, and that I can do so via email, phone, or in person. Add that these exhaust the ways to satisfy my obligation. Since it is obligatory for me to be the first to convey the information to you, the obligation is personal and agential: I am not only responsible for your getting the information, but for delivering it myself. Now it is easy to imagine that although any of the three ways of discharging my obligation is permissible, nonetheless, the more personal the

---

<sup>3</sup> For a quick overview of the theory of normative positions with an explicit agency operator, see [3].

manner of delivery the better morally speaking. Assume also that the three alternatives are mutually exclusive for some reason.<sup>4</sup> We might then describe the options this way. My *doing the minimum* implies my emailing you—this is the way to discharge my obligation in the minimally acceptable way. On the other hand, if I convey the information in person, I will be *doing the maximum* (*what morality recommends*). Finally, we can easily imagine that it is *a matter of moral indifference* that I carry my pen with me when I deliver the message. This illustrates one application of DWE's four primitive operators (tacitly interpreted as relative to some agent, S):

- OB**p: It is *Obligatory* (for S) that p  
**MI**p: The *Minimum* (for S) involves/implies (its being the case that) p  
**MA**p: The *Maximum* (for S) involves/implies (its being the case that) p.  
**IN**p: It is *Indifferent* (for S) that p<sup>5</sup>

We imagine these operators added to some language for classical propositional logic, and taking any formula as argument.<sup>6</sup> Some defined operators, and their intended readings can then be introduced:

- PE**p =<sub>df</sub>  $\sim$ **OB** $\sim$ p. (It is *Permissible* for S that p.)  
**IM**p =<sub>df</sub> **OB** $\sim$ p. (It is *Impermissible* for S that p.)  
**OM**p =<sub>df</sub>  $\sim$ **OB**p. (It is *Omissible* (for S) that p  
**OP**p =<sub>df</sub> **OB**p &  $\sim$ **OB** $\sim$ p. (It is *Optional* for S that p.)  
**SI**p =<sub>df</sub>  $\sim$ **IN**p. (It is *Significant* for S that p.)  
**BC**p =<sub>df</sub> **PE**p & **MI** $\sim$ p. (It is *Beyond the Call* for S that p.)  
**PS**p =<sub>df</sub> **PE**p & **MA** $\sim$ p. (It is *Permissibly Suboptimal* for S that p.)

Continuing our “delicate information” case, note that although the three alternatives (conveying the information by email, phone, or in person) are not on a par morally speaking, each is still *morally optional*. For, each is *permissible*, but none is obligatory for our agent. We saw that doing the minimum involves e-mailing you. What happens if instead I either call or stop by? In each case, I will have done *more than I had to do*—more good than I would have if I had done the minimum permitted. I will have acted “beyond the call of duty”. On the other hand, if I do not stop by, I will not have done what is optimal, but, then if I do still discharge my obligation, it will be done *permissibly, but sub-optimally*, by emailing you or calling you. Furthermore, although each of the three ways of contacting you is optional in either case, I will have done something either beyond the call of duty or I will have done only the minimum; in either case, I will have done something with *moral significance*.

Here is a simple logic cast in this framework<sup>7</sup>, where “\*” ranges over **OB**, **MA**, **MI**:

- A0. All tautologous DWE-wffs;  
A1.  $*(p \rightarrow q) \rightarrow (*p \rightarrow *q)$   
A2. **OB**p  $\rightarrow$  (**MI**p & **MA**p)

<sup>4</sup> We presuppose a simple no-conflicts atmosphere.

<sup>5</sup> Note that the readings are *personal* but not *agential*—a bit more on this below.

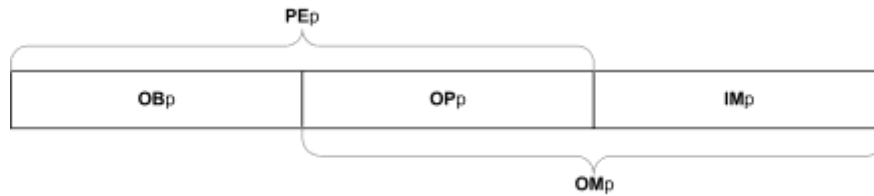
<sup>6</sup> We could easily add an agency operator, BA, to express normative statuses that are both personal and agential, for example. “S *must* bring it about that p” (**OBBA**p) and “S *ought* to bring it about that p” (**MABA**p), “The *least* S can do involves bringing it about that p” (**MIBA**p), etc.

<sup>7</sup> It is easily shown that SDL logics for **OB**, **MA**, and **MI** are derivable from DWE [8], and that completeness can be proven for a semantics like that sketched in the Appendix below [7].

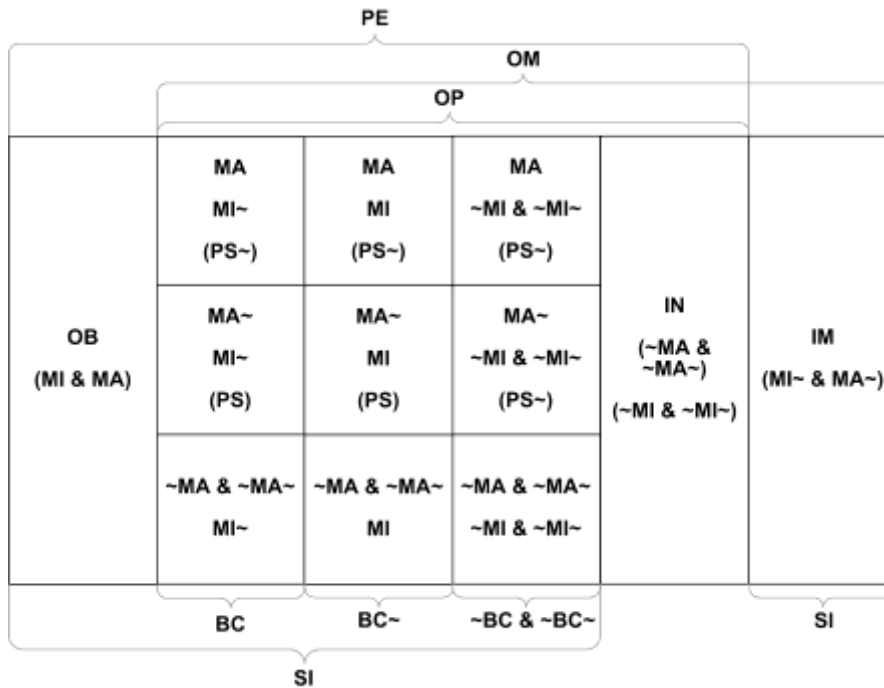
- A3.  $(MIp \vee MAp) \rightarrow PEp$
- A4.  $INp \rightarrow IN\sim p$
- A5.  $INp \rightarrow (\sim MIp \ \& \ \sim MAp)$
- A6.  $(OB(p \rightarrow q) \ \& \ OB(q \rightarrow r) \ \& \ INp \ \& \ INr) \rightarrow INq$

- R1: If  $\vdash p$  and  $\vdash p \rightarrow q$  then  $\vdash q$
- R2: If  $\vdash p$ , then  $\vdash OBp$ .

Recall the traditional three-fold partition of Standard Deontic Logic (SDL):



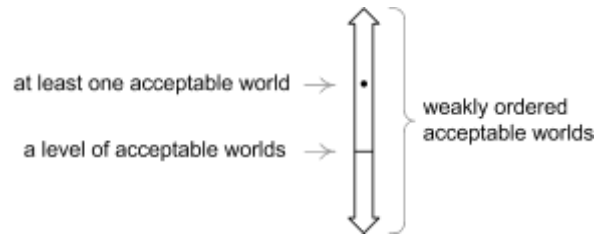
Outer labels (**PE** and **OM**) indicate normative statuses that include more than one normative position. There are thus three normative positions a person may occupy with respect to a proposition from the standpoint of SDL. The increased expressive resources of DWE are reflected in the following analogous partition with twelve normative positions:



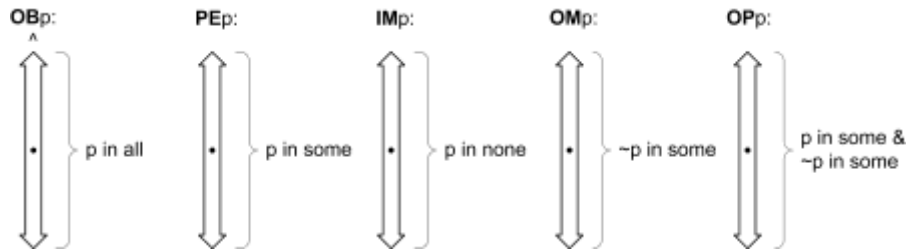
Note that **OP** now subsumes ten normative positions rather than itself being one as with SDL. We turn now to a quick sketch of the semantics.

## 1.2 Semantics

The underlying semantic picture employs a set of worlds, an accessibility relation, and a set of world-relative weak ordering relations. We assume that for any world  $i$ , there is both a set of acceptable alternatives and a morally relevant  $i$ -relative weak ordering of these  $i$ -acceptable alternatives—one that is reflexive, connected, and transitive. Relative to any world  $i$ , we get this picture of the associated structure the semantic framework employs:



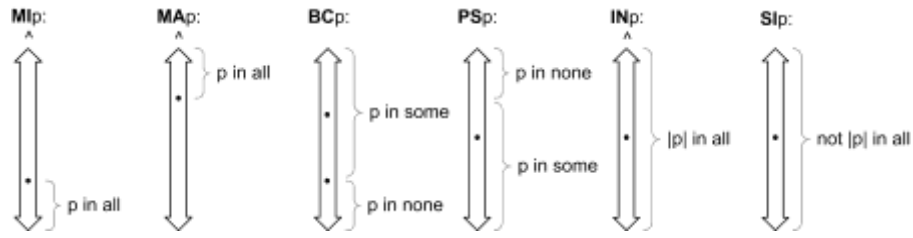
The accessibility relation is interpreted here as relating worlds to their morally *acceptable* alternatives. The dot indicates the context is classical (seriality holds): there is a morally acceptable alternative for every world. The vertical bar represents the weak ordering of these worlds. The horizontal line through the bar represents an equivalence class of  $i$ -acceptable worlds with respect to equi-rank. Note that intuitively, an acceptable alternative need not be one that is *ideal* or *optimal* (though the converse will hold).<sup>8</sup> We do not follow the standard ordering semantics for deontic logic that interprets what is obligatory in terms of what is optimal.<sup>9</sup> Rather, we interpret what is obligatory and what is permissible, etc. in terms of what is acceptable:



(A " $\wedge$ " under an operator indicates that it is primitive in DWE.) So far, the semantics is classical accessibility semantics for the normal modal logic D, and the ordering is only relevant for contrastive purposes. But our other operators will rely on the ordering essentially. As indicated, some acceptable alternatives can outrank others, some can be ranked highest, others lowest among the  $i$ -acceptable worlds, and there may be ties. This allows us to model new operators, including these:

<sup>8</sup> Indeed, the formal semantics for DWE does not assume there need be optimal worlds. It is hard to give a principled defense of such a limit assumption for deontic logics.

<sup>9</sup> Nor alternatively where no limit assumption is assumed, that something is obligatory iff it holds in some optimal range—iff from some point on up, it holds.



The *minimum* involves  $p$  iff the lowest ranked acceptable worlds are  $p$ -worlds; the *maximum* involves  $p$  iff the highest ranked acceptable alternatives are  $p$ -worlds.<sup>10</sup> Thus the minimum and the maximum are mirror images of one another, which results in various symmetries in the logic [8].  $p$  is *beyond the call* iff although there is an acceptable  $p$ -world somewhere, none is among the lowest ranked acceptable alternatives. This guarantees that  $p$  is permissible but precluded by doing the minimum, so it occurs in the acceptable range only above the lower permissible range. Similarly,  $p$  is *permissibly suboptimal* if there is an acceptable  $p$ -alternative somewhere but only below the highest ranked acceptable alternatives. This guarantees that  $p$  is permissible but excluded by doing the maximum. As noted, the ranked acceptable alternatives can be divided into equivalence classes with respect to equal rank—to levels. " $|p|$  in all" above indicates that at each of the associated levels, there is a  $p$ -world as well as a  $\sim p$ -world at that level. If this condition is met by all the  $i$ -acceptable worlds, then  $p$  is deemed a matter of *moral indifference*—there is no level of value among the acceptable alternatives the realization of which hinges on  $p$ 's having a uniform truth value throughout that level. This satisfies "Urmsen's Constraint": what is indifferent is optional but not necessarily the converse [8].  $p$  will be deemed *morally significant* if there is some level of value the realization of which *does* hinge on  $p$ 's or  $\sim p$ 's realization. A formal semantics is given in the Appendix.

### 1.3 A Puzzle: What of the Classical Conception of Supererogation?

DWE was designed to model a richer array of concepts from commonsense morality, and to distinguish concepts typically conflated with one another or neglected in ethical theory and deontic logic. Perhaps one of the most challenging and rich concepts of this sort is that of *supererogation*. It is perhaps now recognized as the classical case of a marginalized or neglected concept in ethical theorizing. Suppose a child is trapped in a burning building. Rescue is dangerous even for trained equipped fire personnel working in teams. A mail-woman on her regular rounds sizes the situation up, and fully recognizing the risks, compassionately enters the building and rescues the child. Her act is supererogatory. There are many real life cases. We all learned in school about Captain Oakes voluntary sacrifice for his comrades and about Florence Nightingale's life of service, and the supererogatory heroism after Flight 90 crashed into the Potomac was viewed worldwide in 1982; but not only heroic sacrifices and service fall under the classical conception. Generally, volunteering, favors, charity, forgiveness, mercy, and tolerance are typically cited are cases of supererogation, and no doubt parents, especially single parents, sometimes make many sacrifices beyond what is required in their dedication to their children. Here is the "The Classical Analysis" of supererogation:

An act is *Supererogatory* for  $S$ : 1) it is optional for  $S$  to do, 2) it is *praiseworthy* for  $S$  to do and 3) it is *not blameworthy* for  $S$  to *not* do.

Condition 1) is the only *deontic condition*: the act is permissible for the agent to perform or to skip. 2) & 3) are "*aretaic*" conditions: conditions that evaluate an agent, evaluating the agent positively for doing the action, but saying no negative evaluation is due if the agent does not perform the action. 1)-3) are certainly necessary on the classical conception of "supererogation" that the analysis is intended to capture.

<sup>10</sup> For ease of exposition, we assume lower and upper limit bounds in the informal glosses.

Now note that DWE is certainly intended to capture the notion of doing more than one has to do, which looks at least like a close cousin of *supererogation*, and the term has been freely (though cautiously) used in [7-11]. Yet there is no mention in DWE of aretaic appraisal at all. No agent-evaluative concepts like praiseworthiness or blameworthiness appear in DWE. Is DWE then missing something vital or even failing altogether to capture one of its key target concepts?

Worse still, the following is often taken for granted:

The Equivalence: An act is Supererogatory iff it is *Beyond the Call* (of duty).

But then if the classical analysis is right, DWE can't claim to be modeling action beyond the call, and indeed, I have already stated that the conditions in the classical analysis are at least necessary on the classical conception of supererogation, so it looks like DWE can't possibly be modeling supererogation, and if the above standard equivalence is right, it can't be modeling action beyond the call either.

Furthermore, in the traditional literature on supererogation, there is a much-discussed mirror image of that concept:

An act is an *Offense\Suberogatory* for S: 1) it is optional for S to do, 2) it is *blameworthy* for S to do and 3) it is not *praiseworthy* for S to not do.

This is also not expressible in DWE, and the closest cousin is that of a permissibly suboptimal option, which is modeled in DWE and is the mirror image operator of BC in DWE. However, as with the classical analysis of supererogation, the only condition expressible in the classical analysis of an offense is condition 1), optionality. So what then is the relationship between an offense and that which we should not do but can (the permissibly suboptimal)? It would seem that that which we can do and shouldn't stands to offenses as action beyond the call does to supererogation, so has DWE failed here to really capture even its avowed concept of permissible suboptimality, much less the classical notion of an offense?

I wish to raise three general questions in this context,

- Q1: What might a preliminary logical framework for aretaic concepts look like?
- Q2: What might such a framework, integrated with a classical deontic logic, look like, one designed to express the *classical analysis* of supererogation (as well as that of offense and related notions)?
- Q3: In particular how does such a framework fit in with DWE's account of *doing more good than one must*, and *permissibly doing less than one ought*, which invoke no aretaic notions at all?

I will focus here primarily on Q1 and Q2, and end by sketching an answer to Q3 to show a) that the classical *analysis* of supererogation is not adequate to its own *conception* of supererogatory acts, though never noticed before, and b) that the concepts modeled in DWE used to show this lack in the classical analysis are more fundamental. The context will be classical throughout (e.g. conflicts of normative appraisal excluded).

I begin sketching a Kangerian-Andersonian framework for predetermination and obligation, with a person-relative intended interpretation of each.

## 2 A Modal Framework for Predetermination and Obligation

The main operator in our framework for predetermination is just an interpreted classical necessity operator, PRp, with this intended interpretation: "It is (as of now) predetermined (for Jane Doe) that p". We use standard Kripke

structures consisting of a set of worlds,  $W$ , with a relation,  $CO \subseteq W \times W$ . We intend that  $COij$  iff what happens in  $j$  is consistent with our agent's current abilities and disabilities at  $i$ . The truth condition for PR is the usual one:  $M \models_i PRp$  iff  $\forall j(COij \rightarrow M \models_j p)$ . We introduce the dual, "it is consistent with our agent's abilities that  $p$ ":  $COP =_{df} \sim PR \sim p$ , and its derived truth condition:  $M \models_i COP$  iff  $\exists j(COij \ \& \ \Box \models_i p)$ . We add a single constraint:  $CO\text{-RFLX}$ :  $COii$ . The worlds consistent with our agent's abilities at a given world,  $i$ , might then be thought of as the *i-accessible worlds*:  $CO^i = \{j: \langle i, j \rangle \in CO\}$ . It will also prove convenient to introduce a notation for *the set of all propositions consistent with our agent's abilities*:  $CO_i = \{X: X \cap CO^i \neq \emptyset\}$ .  $CO^i$  contains every world consistent with our agent's abilities at  $i$ , whereas  $CO_i$  contains the set of propositions true at some such world. Note that the existence of a  $p$ -world *consistent with* my abilities does not entail that  $p$  is *within* my abilities. Just consider any tautology, or any independent action someone else may or may not perform. But it does mean that such a  $p$ -world cannot involve me having or exercising some ability I lack. The well known normal modal logic, KT, for PR (PR-KT) is determined by the class of  $CO$ -reflexive models.

We now add an Andersonian-Kangerian constant,  $d$ , for "The demands on Jane Doe are all met" (or "Jane Doe's responsibilities are all met").<sup>11</sup> We represent the extension of " $d$ " as a set of worlds,  $DEM$ ,  $DEM \subseteq W$ , and we give " $d$ "'s truth-conditions accordingly:  $M \models_i d$  iff  $i \in DEM$ . We define a *non-agential but personal obligation operator*:

$$OBp =_{df} PR(d \rightarrow p),$$

and read it as follows:  $OBp$  iff it is obligatory *for Jane Doe* to be such that  $p$ .<sup>12</sup> We add one axiom to PR-KT governing our deontic constant:  $\vdash COD$  (i.e.  $\sim PR \sim d$ ). " $COD$ " says  $d$ 's truth is *consistent with* Jane Doe's abilities, but it does not say it is *within* her abilities, for good reason.<sup>13</sup> Axiom  $d$  is validated by the condition that satisfying Doe's responsibilities is consistent with her abilities:  $\forall i \exists j(COij \ \& \ j \in DEM)$ . Call the resulting system "PR-KTd". As is well known, it is characterized by the class of all models satisfying this constraint ([11]).

SDL,

SL:	All Tautologies
OB-K:	$OB(p \rightarrow q) \rightarrow (OBp \rightarrow OBq)$
OB-NC:	$OBp \rightarrow \sim OB \sim p$
MP:	If $\vdash p$ and $\vdash p \rightarrow q$ then $\vdash q$
OB-NEC:	If $\vdash p$ then $\vdash OBp$ ,

is part of the pure deontic fragment of PR-KTd.<sup>14</sup>

<sup>11</sup> So this constant is agent relative by intention. With a framework for multiple agents, we would need either multiple constants, or another way to individualize and differentiate the distinct demands morality places on each of us given our circumstances, social relationships, past commitments, etc., as well as a way to tie these into our distinct abilities/disabilities, if we want to retain Kant's Law.

<sup>12</sup> The intended reading of "**OB**" is developed and defended in [13]. It doesn't express the *impersonal* notion *it is obligatory that p*. It expresses a *personal* obligation our agent is under. Nonetheless, it does not require that she be *the agent of p*. We take the form of a personal obligation as an obligation *to be such that p*, and we then take an agential obligation to be a special case of a personal obligation, one to the effect that our agent has *to be such that she herself brings it about that p*, and thus to be a compound of a personal obligation operator and an agency operator. We pass over agency and agential obligations here and allow the personal obligation operator and our person-relative modal notions to serve.

<sup>13</sup> Jane Doe may have delegated the last step in *her* project to her assistant, and it may now be predetermined for Doe that her project will be completed only if the assistant completes it, which she will. The project's completion is no longer within Doe's ability, but it is still consistent with her ability. Now just add that the project's completion is equivalent to  $d$ . For more on the distinction, see [13].

<sup>14</sup> In fact the stronger system that results from adding " $OB(OBp \rightarrow p)$ " to SDL corresponds to KTd. See [11].



Plainly we are engaged in considerable idealization, but this simple familiar system allows us to take some first steps toward a more comprehensive integration of aretaic notions with deontic ones.

### 3 A Preliminary *Framework for Aretaic Appraisal (FAA)* of Agents

#### 3.1 Aretaic Preference and Aretaic Appraisal

Some states of affairs reflect favorably on people, others unfavorably, some more favorably than others, and some neutrally. I sketch a simple framework here that allows for such agent appraisals, simplifying and slightly modifying that in [12], which gives more details. We stick to all things considered agent appraisal throughout.

We first define a world-relative ordering function, which will yield a weak or quasi-ordering relation,  $\succeq_i$ :

$$\succeq: W \rightarrow \mathcal{P}(\mathcal{P}(W) \times \mathcal{P}(W)) \text{ [i.e. } \succeq_i \subseteq \mathcal{P}(W) \times \mathcal{P}(W)\text{].}$$

For each world  $i$ , and proposition pair,  $X$  and  $Y$ ,  $X \succeq_i Y$  if and only if  $X$  *reflects as well on our agent as*  $Y$  ( $X$  is *aretaically as good as*  $Y$ ) *from the standpoint of*  $i$ . We introduce a corresponding binary operator to take any pair of formulas:

$$M \models_i p \geq q: \|p\|^M \succeq_i \|q\|^M.$$

We evaluate agents for their actions, results of their actions, motives for acting, intentions in acting, traits of character, etc.<sup>15</sup> So as to allow for this variety, the relata of our ordering relation excludes only propositions inconsistent with our agent's abilities:

$$\succeq\text{-}CO_i \text{ Confinement: } \forall i (\succeq_i \subseteq CO_i \times CO_i).$$

We will also assume that all propositions consistent with our agent's abilities are self-comparable, and we will assume transitivity as well

$$\begin{aligned} \text{Reflexive: } & \forall i \forall X (X \in CO_i \rightarrow X \succeq_i X) \\ \text{Transitive: } & \forall i \forall X \forall Y \forall Z [(X \succeq_i Y \ \& \ Y \succeq_i Z) \rightarrow X \succeq_i Z].^{16} \end{aligned}$$

We do *not* endorse  $\succeq$ -Connectivity,  $\forall i \forall X \forall Y [X, Y \in CO_i \rightarrow (X \succeq_i Y \vee Y \succeq_i X)]$  for FAA as a basic constraint. It is not obvious that *any* two propositions consistent with our agent's abilities must be aretaically comparable since they may involve very different grounds for praise or blame. However, we will need it later to explore supererogation. We consider further constraints below, in the context of discussing the concepts of neutral, positive and negative aretaic appraisal of an agent.

The following basic schemata and rules are validated for FAA, so take them here as axiomatic:

$$\begin{aligned} \succeq\text{-}CO \text{ Confinement: } & \vdash p \geq q \rightarrow (CO_p \ \& \ CO_q) \\ CO\text{-Rflx}(\succeq): & \vdash CO_p \rightarrow p \geq p \end{aligned}$$

<sup>15</sup> We assume here that evaluating an agent for a state of affairs allows us to derivatively evaluate an agent for these other things derivatively.

<sup>16</sup> Clearly, confinement and reflexivity imply  $\forall i \forall X (X \in CO_i \leftrightarrow \exists Y (X \succeq_i Y \vee Y \succeq_i X))$ . We might thus designate the *aretaically evaluable* propositions, as those comparable with some proposition or other, or those self-comparable. In turn, we might designate the propositions consistent with our agent's abilities as those that are aretaically evaluable.

Trans( $\geq$ ):	$\vdash (p \geq q \ \& \ q \geq r) \rightarrow p \geq r$
$\geq$ -RE1:	If $\vdash p \leftrightarrow q$ then $\vdash r \geq p \rightarrow r \geq q$
$\geq$ -RE2:	If $\vdash p \leftrightarrow q$ then $\vdash p \geq r \rightarrow q \geq r$ .

A *strong preference* relation and an *equi-ranking* relation are definable in familiar ways:  $X >_i Y =_{df} X \geq_i Y \ \& \ \sim(Y \geq_i X)$ ,  $X =_i Y =_{df} X \geq_i Y \ \& \ Y \geq_i X$ . Similarly for the corresponding operators:  $p > q =_{df} p \geq q \ \& \ \sim(q \geq p)$  and  $p \approx q =_{df} p \geq q \ \& \ q \geq p$ . Derivative truth-conditions for these operators are:  $M \models_i p > q: \|p\|^M > \|q\|^M$  and  $M \models_i p \approx q: \|p\|^M = \|q\|^M$ . From these axioms, rules, and our definitions, various familiar properties for  $>$  and  $\approx$  are derivable, as well as these linking them to CO:  $COp \leftrightarrow p \geq p$ ;  $COp \rightarrow p \approx p$ ;  $p \approx q \rightarrow (COp \ \& \ COq)$ ; and  $p > q \rightarrow (COp \ \& \ COq)$ .<sup>17</sup>

### 3.2 Neutrality

We take a neutral proposition to be one consistent with an agent's abilities but reflecting no overall positive or negative merit all-in-all on our agent, perhaps because it involves no positive or negative aretaic components at all or because it has an equal balance of positive and negative aretaic value, "neutralizing" the two opposing values. We treat tautological propositions as "anchors" in this scheme: they reflect neutrally on agents, if anything can, so we take this to be a matter of logic. We then define our first notion, "aretaic neutrality", relying on these anchors--"T" for the verum (always true):

$$ANp =_{df} p \approx T,$$

and given this definition, ANT plainly follows from the preceding.

Since we are especially interested in the positive and negative aretaic appraisal of things consistent with our agent's ability, paving the way for linking such appraisal with the deontic appraisal of propositions consistent with our agent's abilities, we want, and get (it is now derivable from the above axioms and rules):

$$AN-CO: ANp \rightarrow COp,$$

as well as  $\sim AN\perp$  and  $ANp \rightarrow (ANq \leftrightarrow p \approx q)$ , " $\perp$ " for the falsum (always false). This AN-RE rule follows readily from the earlier RE principles for  $\geq$  and our definitions: If  $\vdash p \leftrightarrow q$  then  $\vdash ANp \leftrightarrow ANq$ .

Note however that neutrality is not indifferent to negation,  $\not\models ANp \leftrightarrow AN\sim p$ , and this would entail  $AN\perp$ , given ANT, and thus  $CO\perp$ . But even where  $COp$  and  $CO\sim p$ , neutrality is still not intuitively indifferent to negation: that *I do not bring it about that I now do some wonderful thing* might be consistent with my ability, as might its negation, and it might very well be neutral (e.g. there is nothing special about this opportunity to do good), but that *I do bring about something wonderful right now* (p) may not be neutral.

### 3.3 Indifference

Some propositions will be *aretaically indifferent* for our imagined agent.—whether true or false they will not reflect positively or negatively on her. We define this notion as:

$$AIp =_{df} ANp \ \& \ AN\sim p.$$

As stated earlier, aretaic indifference should be stronger than mere neutrality, and by definition,  $AIp \leftrightarrow (p \approx T \ \& \ \sim p \approx T)$ , so given RE for  $\approx$ , we get the mark of a true indifference notion:  $AIp \leftrightarrow AI\sim p$ . Also derivable are:  $AIp \rightarrow ANp$ ,  $AIp \ \& \ AIq \rightarrow p \approx q$ ,  $AIp \rightarrow p \approx \sim p$ ,  $\sim AI\perp$ ;  $\sim AI\perp$ ,  $AIp \rightarrow COp$ , and this AI-RE rule: if  $\vdash p \leftrightarrow q$  then  $\vdash AIp \leftrightarrow AIq$ .

<sup>17</sup> More nuanced and careful exploration of the logic of  $\geq$  will have to await another occasion.

### 3.4 Praiseworthiness and Blameworthiness

We take the praiseworthy (blameworthy) propositions as those ranked aretaically higher (lower) than neutral propositions, and this idea is captured by these concise definitions:

$$PWp \text{ =}_{df} p > \top \text{ and } BWp \text{ =}_{df} \top > p.$$

The derivative truth conditions are:  $M \models_i PWp$  iff  $\|p\|^M >_i W$ ;  $M \models_i BWp$  iff  $W >_i \|p\|^M$ . The following principles are validated and now derivable given these definitions:

$$\begin{aligned} PW/BW\text{-CO:} & (PWp \vee BWp) \rightarrow COp \\ PW\text{-BW EXCL:} & PWp \rightarrow \sim BWp \\ PW\text{-AN/AI EXCL:} & PWp \rightarrow \sim (ANp \vee AIp) \\ BW\text{-AN/AI EXCL:} & BWp \rightarrow \sim (ANp \vee AIp) \\ PW > AN > BW: & (PWp \ \& \ ANq \ \& \ BWr) \rightarrow (p > q \ \& \ q > r) \end{aligned}$$

The following indifference exclusion principle is also derivable:

$$AI\text{-EXCL: } AIp \rightarrow (COp \ \& \ \sim BWp \ \& \ \sim BW\sim p \ \& \ \sim PWp \ \& \ \sim PW\sim p).$$

The only thing that blocks the converse of AI-EXCL,

$$AI\text{-EXCL': } (COp \ \& \ \sim BWp \ \& \ \sim BW\sim p \ \& \ \sim PWp \ \& \ \sim PW\sim p) \rightarrow AIp,$$

is incomparability. There may be propositions consistent with our agent's ability (and thus each comparable to itself), that are nonetheless not comparable to  $\top$ , and so not "placed" above, below, or among the neutrals. Thus  $p$  might satisfy the left side of AI-EXCL' merely because it is incomparable with  $\top$ . Such propositions would presumably contain conflicting positive and negative aretaic components separately pulling above and below the neutral line in a way that doesn't allow for resolution. Likewise, we don't have:  $COp \rightarrow (ANp \vee PWp \vee BWp)$ . If, however, we add

$$\geq\text{-Connectivity: } \forall i \forall X \forall Y [(X, Y \in CO_i) \rightarrow (X \geq_i Y \vee Y \geq_i X)],$$

to FAA--call the result  $FAA^C$ , we validate a comparability axiom for this strengthened framework,

$$CO\text{-COMP: } \vdash (COp \ \& \ COq) \rightarrow (p \geq q \vee q \geq p),$$

and then the following are all validated and derivable:

$$\begin{aligned} CO\text{-COMP':} & (COp \ \& \ COq) \rightarrow (p > q \vee q > p \vee p \approx q) \\ CO\text{-DEF'':} & COp \leftrightarrow (p \geq \top \vee \top \geq p) \\ & COp \leftrightarrow (p > \top \vee \top > p \vee p \approx \top) \\ AN\text{-DEF':} & ANp \leftrightarrow (COp \ \& \ \sim BWp \ \& \ \sim PWp) \\ AI\text{-EXCL':} & (COp \ \& \ \sim BWp \ \& \ \sim BW\sim p \ \& \ \sim PWp \ \& \ \sim PW\sim p) \rightarrow AIp \\ AR\text{-EXH:} & COp \leftrightarrow (ANp \vee PWp \vee BWp) \\ AI\text{-DEF':} & AIp \leftrightarrow (COp \ \& \ \sim BWp \ \& \ \sim BW\sim p \ \& \ \sim PWp \ \& \ \sim PW\sim p). \end{aligned}$$

AI-DEF is essential to the classical framework for supererogation that it is one of our central aims to explore here, so we will assume it henceforth.<sup>18</sup>

Do praiseworthiness and blameworthiness satisfy no conflicts principles:

$$\begin{aligned} \text{PW-NC: } & \text{PW}p \rightarrow \sim\text{PW}\sim p \\ \text{BW-NC: } & \text{BW}p \rightarrow \sim\text{BW}\sim p? \end{aligned}$$

These seem plausible for "all-in-all" readings. For suppose you would be praiseworthy (all in all) for being kind or for saving the drowning child. It does not seem right to say that it is also possible that you would be praiseworthy (all in all) for the negation of these very things. Also, PW-NC and BW-NC are clearly presupposed in the classical conceptions of supererogation and offense. So we add these constraints to get "FAA<sup>CNC</sup>":

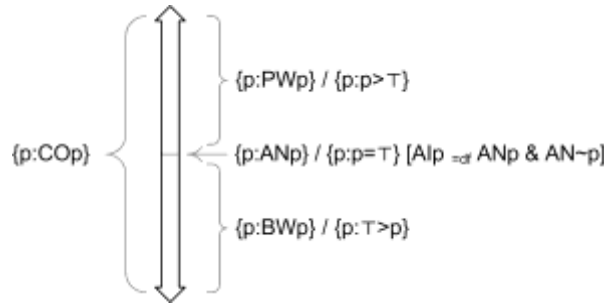
$$\begin{aligned} \text{PW-NC': } & \forall i \forall X (X >_i W \rightarrow \sim[(W-X) >_i W]) \\ \text{BW-NC': } & \forall i \forall X (W >_i X \rightarrow \sim[W >_i (W-X)]) \end{aligned}$$

PW-NC' tells us that for any world  $i$ , and proposition  $X$ , if  $X$  is ranked higher than the tautological proposition, then the negation of  $X$  is not also ranked higher than that proposition. BW-NC' gives the mirror image. These two validate the following *upper and lower exclusion* principles, to be taken as axiomatic for FAA<sup>CNC</sup>:

$$\begin{aligned} \text{T} > \text{EXCL: } & \vdash p > \text{T} \rightarrow \sim(\sim p > \text{T}) \\ \text{T} < \text{EXCL: } & \vdash \text{T} > p \rightarrow \sim(\text{T} > \sim p).^{19} \end{aligned}$$

PW-NC and BW-NC are then derivable by definition.

We will presuppose the stronger FAA<sup>CNC</sup> henceforth, so that we are operating with these stronger principles generated by connectivity and no conflict semantic constraints added to FAA, as they are essential for understanding the standard conception of supererogation and they will facilitate our simple applications to show the fruitfulness and potential of a mixed deontic-aretaic scheme.<sup>20</sup> Here is a simple picture for the FAA<sup>CNC</sup> framework:



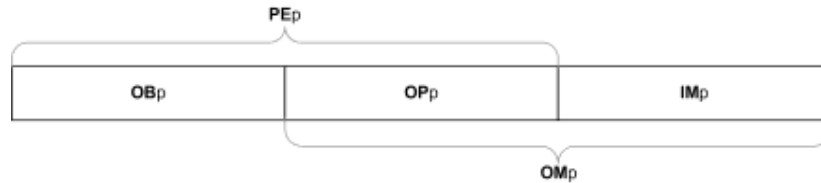
<sup>18</sup> We could add a weaker semantic assumption to validate just the last two: W->Connectivity:  $\forall i \forall X [(X \in CO_i) \rightarrow (X \geq_i W \vee W \geq_i X)]$ , but it is hard to see any principled reason to accept comparability to  $W$  for each  $CO_i$  proposition, but not full comparability of all such propositions to one another.

<sup>19</sup> Generalized, we would have  $p > q \rightarrow \sim(\sim q > p)$ : intuitively, if  $p$  reflects better on an agent all in all than  $q$ , then the negation of  $q$  cannot reflect better on the agent all in all than  $p$ . Similarly for the corresponding generalizations of the semantic constraints, PW-NC' and BW-NC'.

<sup>20</sup> Obviously we could add each no conflicts principle separately (although it is unclear if it is very plausible to have one and not the other), and we could add both to FAA alone to get FAA<sup>NC</sup> without comparability, which is worth exploring, but FAA<sup>CNC</sup> is most relevant to the applications to follow.

#### 4 Aretaic Partitions & Aretaic Normative Positions

It is well known that in the traditional deontic systems, all propositions are partitioned into three mutually exclusive and exhaustive classes, those *obligatory*, those *impermissible* and those *optional* (often mislabeled "indifferent"):

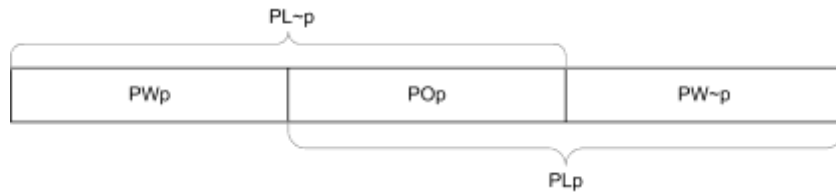


Similar relationships hold for PW and BW, but first let us introduce some analogous abbreviations:

$PLp =_{df} \sim PWp$ :	It is Praise-Less that p.
$POp =_{df} \sim PWp \ \& \ \sim PW\sim p$ :	It is Praise Optional that p
$BLp =_{df} \sim BWp$ :	It is Blame-Less that p
$BOp =_{df} \sim BWp \ \& \ \sim BW\sim p$ :	It is Blame-Optional that p

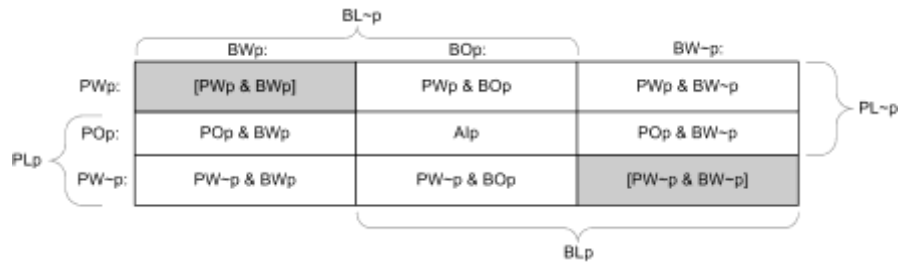
Given COMP, the following hold,  $AIp \leftrightarrow (COp \ \& \ BLp \ \& \ BL\sim p \ \& \ PLp \ \& \ PL\sim p)$ ;  $AIp \leftrightarrow (COp \ \& \ BOp \ \& \ POp)$ ; and  $COp \rightarrow [AIp \leftrightarrow (BOp \ \& \ POp)]$ .

For both PW and BW we get exactly analogous partitions. Here is the PW-Partition (PW-P):



PW-P is this conjunction: a)  $PWp \vee PW\sim p \vee (\sim PWp \ \& \ \sim PW\sim p)$  & b)  $\sim(PWp \ \& \ PW\sim p)$  & c)  $\sim(PWp \ \& \ (\sim PWp \ \& \ \sim PW\sim p))$  & d)  $\sim(PW\sim p \ \& \ (\sim PWp \ \& \ \sim PW\sim p))$ . Clearly, a), c) and d) are just truth-functional tautologies. Only b) is not, and it is just No PW Conflicts again, PW-NC:  $PWp \rightarrow \sim PW\sim p$ . So PW-P is equivalent to PW-NC: PW-P iff PW-NC. The BW Partition (BW-P) is perfectly analogous and similarly reduces to No BW Conflicts principle, so BW-P iff BW-NC.

What happens when we consider compounding these two aretaic partitions and classify options in terms of *both* the positive and the negative aretaic operators above?



Nine possible combinations are indicated, of which there are seven new aretaic normative positions. The two eliminations in shaded boxes follow from our earlier theorem,  $PWp \rightarrow \sim BWp$ , which derives from our definitions and the thesis,  $p > T \rightarrow \sim(T > p)$ . Furthermore, the standard conception of supererogation and offense presuppose the exclusiveness of all-in-all praiseworthiness and blameworthiness, for else an act that was supererogatory and thus praiseworthy to do might nonetheless be blameworthy to do, which jars. Call the resulting 7-fold Partition "PW-BW P".<sup>21</sup> We turn now to integrating this aretaic framework with a standard deontic one.

### 5 Aretaic-Deontic Partitions and Some Underlying Issues

As we noted earlier, SDL entails an OB-Partition. What happens when we combine the three deontic categories with the preceding seven aretaic ones? Ignoring the shading and the text in parentheses and brackets for now, we get this 21-fold partition:

	PWp & BOp:	PWp & BW~p:	POp & BWp:	Alp:	POp & BW~p:	PW~p & BWp:	PW~p & BOp:
OBp:	PWp & BOp & <b>OBp</b> [elim by f]	PWp & BW~p & <b>OBp</b>	POp & BWp & <b>OBp</b> [elim by b/d/e/f]	Alp & <b>OBp</b> [elim by e/f]	POp & BW~p & <b>OBp</b> [elim by e]	PW~p & BWp & <b>OBp</b> [elim by a - f]	PW~p & BOp & <b>OBp</b> [elim by a/c/e]
OPp:	PWp & BOp & <b>OPp</b> (SU <sup>2</sup> p) [elim by c]	PWp & BW~p & <b>OPp</b> (QSp) [elim by c/d]	POp & BWp & <b>OPp</b> (OF <sup>2</sup> p) [elim by d]	Alp & <b>OPp</b> (Fip)	POp & BW~p & <b>OPp</b> (OF <sup>2</sup> ~p) [elim by d]	PW~p & BWp & <b>OPp</b> (QOp) [elim by c/d]	PW~p & BOp & <b>OPp</b> (SU <sup>2</sup> ~p) [elim by c]
IMp:	PWp & BOp & <b>IMp</b> [elim by a/c/f]	PWp & BW~p & <b>IMp</b> [elim by a - f]	POp & BWp & <b>IMp</b> [elim by e]	Alp & <b>IMp</b> [elim by e/f]	POp & BW~p & <b>IMp</b> [elim by b/d/e/f]	PW~p & BWp & <b>IMp</b>	PW~p & BOp & <b>IMp</b> [elim by f]

We have already eliminated (BWp & PWp) and (BW~p & PW~p) in our framework, so there are no labels for those combinations. On top of the seven columns we have the seven cell labels from the aretaic partition, and left of the three rows, we have the prior deontic cell labels. As with the 7-fold partition, the 21-fold partition inherits the exhaustiveness and exclusiveness of the parent partitions, OB-P and PW-BW P.

<sup>21</sup>The partition inherits the mutual exclusiveness and exhaustiveness of the two three-fold schemes that generated it. The exhaustiveness of the BW-P partition entails that if PWp, then p must satisfy that label as well as one of the three column labels, and thus find a place in at least one box in the top row. Similarly for the supposition if POp, or if PW~p. But now by the exhaustiveness of the PW-P partition, every p must satisfy at least one of these three antecedent conditions—it must satisfy one of the three row labels. So by a 3-part version of Constructive Dilemma, every p must fit into one of the nine boxes, and since no p fits into the two red boxes in our framework, it follows that every p fits into at least one of the 7 white boxes. Similar reasoning about inheritance will show that no p can satisfy more than one of the labels in the boxes, for that would be inconsistent with the non-exclusiveness of the parent partitions, PW-P and BW-P.

Note that we can define and identify a variety of moral concepts of interest in this framework. *Culpable obligations*—obligations one would be blameworthy to omit ( $=_{df}$  OBp & BW~p) appear in the top row second and fifth columns, and given  $IMP =_{df} OB\sim p$ , in the third and sixth columns of the third row. Similarly for non-culpable obligations ( $=_{df}$  OBp & BL~p). We can then raise interesting questions such as "Can there be obligations of either of these kinds?", or even "Can there be obligations that are *blameworthy* for Jane Doe to *fulfill* or *praiseworthy* to *violate*?", and we can identify how positive answers would fit in the above scheme and explore the eliminative implications of negative answers. Which normative positions or statuses can be expressed and instantiated is of fundamental importance in ethical theorizing, so this is a place where the systematicity of deontic logic has a greater chance of being of genuine aid. We will illustrate this in the remainder of this section and the next by considering some simple theses connecting aretaic and deontic concepts and their impact on the above partition.

Many would endorse two basic bridging theses at first glance:

a) No IM-PW Conflicts:  $PWp \rightarrow PEp$

b) No OB-BW Conflicts:  $BWp \rightarrow PE\sim p$ .

These are reductive theses, since it is easy to see that they eliminate the realization of certain normative positions. The six respective eliminations (three each) that these entail are indicated in the darkest shaded boxes in the top and bottom rows. (There is no impact on the middle row.) The result would be a reduction of deontic-aretaic statuses to a 15-fold partition. However, having illustrated the reductions a) and b) imply, and despite the fact that these are often taken for granted by friends of supererogation and ethical theorists generally, there are reasons to be less sure on reflection.

*Doubts about a):* First note that it is widely thought that we lack obligation omniscience: one can have an obligation that p and not realize one has that obligation, even in the case of a moral obligation. For example, I may have forgotten to pay you back \$20 that I borrowed, where no other morally relevant considerations come to bear other than the promise to pay you back. Similarly, I can be non-culpably ignorant of certain facts and as a result not realize that p is impermissible—for example, spending \$20 on something quite optional when I then can't pay you back on time as a result. Now add that for some such p, were it not for the facts of which I am ignorant, it would be very good that p, and that I bring about p motivated by just such a belief about p's goodness. It seems that in this case, I am praiseworthy, all things considered, for p even though p is impermissible. For example, suppose I give a very substantial sum to a charity for stray cats shortly after, unbeknownst to me, my savings have been lost in a stock market crash. As a result, I'm obligated to not give to the charity, since my children will need every penny I have—assume I'm a single parent; but given my blameless ignorance at the moment, and my very good intentions at the time, I am all things considered praiseworthy for giving to the charity, even though doing so was not permissible, unbeknownst to me.

*Doubts about b):* Similarly, suppose again that I am subject to the same non-culpable ignorance of my sudden loss of savings and, so unbeknownst to me, it is obligatory for me to hang on to every cent I have for my family. Now add that I could help out a good friend who has helped me before by giving her \$20, and as far as I know, I could do this permissibly, and at trivial cost. Yet I refuse to do so for the most selfish and callous of reasons. It then seems that all things considered I am blameworthy for not helping my friend in these circumstances, despite the fact that, unbeknownst to me, my familial obligations make it overridingly obligatory to not help. Lest there be doubts, consider that a friend in the know clearly can aptly say: "Paul you *can't* do that! The stock market just crashed. If you do that your kids will starve! You *must* hold onto

what you have..." The "can't" here is plausibly a “can't” of moral impermissibility and the “must” a “must” of deontic necessity (see [9]), but the claim would have to be false if I could not have an obligation to not give to the charity until I realized I had it or realized its conclusive grounds. It seems blameworthiness must be sharply separated from deontic necessity.<sup>22</sup>

Often friends of supererogation tacitly endorse a) and b) in the way they define such acts, but we will be more cautious, as it is questionable that they are deontic-logical truths. I turn to supererogation and kin now.

## 6 These Partitions and the Classical Analysis of Supererogation

The *classical analyses* of supererogation and of offense (suberogation) are:

$$\begin{aligned} \text{SU}^{\text{a}}\text{p}: & \text{PWp} \ \& \ \sim\text{BW}\sim\text{p} \ \& \ \text{OPp} \\ \text{OF}^{\text{a}}\text{p}: & \ \text{BWp} \ \& \ \sim\text{PW}\sim\text{p} \ \& \ \text{OPp} \end{aligned}$$

Something is then supererogatory (for Jane) iff it is praiseworthy, its negation is not blameworthy and it is (deontically) optional. In contrast something is an offense (suberogatory) if it is blameworthy, its negation is not praiseworthy, and it is optional.

[15] and [16] propose extending the classical scheme by adding acts of "quasi-supererogation" and "quasi-offense", and argue for their possible realizations:

$$\begin{aligned} \text{QSp}: & \ \text{PWp} \ \& \ \text{BW}\sim\text{p} \ \& \ \text{OPp} \\ \text{QOp}: & \ \text{BWp} \ \& \ \text{PW}\sim\text{p} \ \& \ \text{OPp} \end{aligned}$$

Something is quasi-supererogatory (for Jane) iff it is praiseworthy, its negation *is* blameworthy and it is optional; something is a quasi-offense (quasi-suberogatory) if it is blameworthy, its negation *is* praiseworthy, and it is optional.

Let us introduce only one more mixed concept--FI, for "Full Indifference":

$$\text{FIp} \text{ =}_{\text{df}} \text{OPp} \ \& \ \text{AIp}.$$

Note that stronger concepts of indifference are clearly possible with a stronger deontic component like that in DWE (to wit: INp & AIp). So the "full" here is purely contextual—as full as it can get relying only on SDL's and FAA<sup>CNC</sup>'s resources. Blending DWE's deontic concepts with the aretaic ones above must wait for another occasion.

These five new concepts are easily accommodated and are present already in our prior 21-fold partition. Since they entail deontic optionality they occur only in the middle row of that partition and are indicated in parenthesis near the bottom of each cell in that row. The result suggests that lingering behind the classical conception of supererogation is a framework with 21 potential categories, far more than previously articulated.<sup>23</sup> Some of their logical features are also revealed at a glance, for example that the five new operators are mutually exclusive and that if something is optional and not fully indifferent then it will satisfy one of the first four operators, and as alluded to earlier when we endorsed PW-BW Exclusion, no supererogatory or quasi-supererogatory option is blameworthy, and similarly, no offense or quasi-offense is praiseworthy.<sup>24</sup>

<sup>22</sup> Although beyond the scope of the current paper, the two principles above perhaps look plausible at first glance because we tend to conflate them with the genuinely plausible principles we get if we replace the partially agent-evaluative notions of blameworthiness and praiseworthiness with the more purely state-of-affairs-evaluative notions of *goodness* and *badness*. When we call acts or results of acts “good” or “permissible”, we evaluate them independently of our evaluation of the agent’s motives, etc., so it is plausible to expect stronger links here. I explore this elsewhere.

<sup>23</sup> [16] identifies nine exclusive statuses, and unlike here, that is nine dependent partially on introducing action concepts in the scope of operators, but he also indicates he makes no claim the scheme is complete.

<sup>24</sup> I know of no friend of supererogation that ever felt the need to add to the definition of supererogation that it was also not blameworthy to do, and similarly for the definition of offense with respect to praiseworthiness. There is a strong



In the late twentieth century, supererogation was a hotly contested concept, with many arguing against its existence. For example, [18] argues for the rejection of supererogation by roughly endorsing the following aretaic-deontic bridging principle:

c)  $PWp \rightarrow OBp$ .

Clearly if we add this scheme, we get  $\vdash \sim(SU^a p \vee QSp \vee QOp)$ . Only offenses remain. Note c) also entails a)  $PWp \rightarrow PEp$ , given  $OBp \rightarrow PEp$ , so a)'s eliminations would follow from c) as well. Furthermore, it is highly unlikely that anyone accepting c) would not also endorse another "if strong aretaic then strong deontic" bridging principle:

d)  $BWp \rightarrow IMPp$ .

Indeed, I believe that d) is more widely endorsed than c). If we add this scheme, it yields  $\vdash \sim(OF^a p \vee QSp \vee QOp)$ , and since d) entails our earlier b)  $BWp \rightarrow PE\sim p$ , b)'s prior eliminations follow as well. So under either c) or d), the quasi-notions are eliminated. If both c) and d) are endorsed, the whole middle row save the central category of full indifference is eliminated, as well as those positions shaded darkest in the top and bottom row, leaving us with just nine possible positions: the three white ones and six lightest shaded ones. This reduction of the middle row to one cell is one version of what I have called Moral Rigor (MR):  $\vdash OPp \leftrightarrow FIp$ . The highly contentious direction is left to right, that everything neither obligatory nor permissible reflects indifferently on an agent. Correlatively, the eliminations in the partition also reflect a version of Strong Exhaustion (SEX), that every option is either obligatory, impermissible or (fully) indifferent:  $\vdash OBp \vee IMPp \vee FIp$ . It is a far cry in its substantive import from the Traditional Exhaustion formula:  $\vdash OBp \vee IMPp \vee OPp$ . MR and SEX rule out not only supererogation but all non-indifferent optional normative positions. If one thinks of the options that are neither obligatory nor impermissible as *totally indifferent*, one is, perhaps unwittingly, ruling out all the non-indifferent *options* above. Lastly, let me note two other theses in the neighborhood, the converses of c) and d):

e)  $OBp \rightarrow PWp$

f)  $IMPp \rightarrow BWp$ ,

asserting "if strong deontic then strong aretaic" bridging principles.<sup>25</sup> The new eliminations that ensue solely from these are indicated in the lightest shaded cells in the top and bottom rows (some prior eliminations follow from e) and f) too, as indicated in square brackets in the same rows). They do not impact the middle optionality row, and so do not impact the traditional optionality-entailing concepts of supererogation and kin defined above. However, if we conjoin c)-f), that is, change c) and d) into *bi*-conditionals, then since c) and d) entail a) and b) as we have already seen, the result is that all that is left are the three unshaded cells. That is, if c)-f) are jointly assumed, then we get this thesis (" $\nabla$ " for an exclusive-or operator):

$(OBp \ \& \ PWp \ \& \ BW\sim p) \ \nabla \ FIp \ \nabla \ (IMPp \ \& \ PW\sim p \ \& \ BWp)$ .

---

presupposition in favor of PW-BW Exclusion in the ethical theory tradition focused on these notions, and probably in general.

<sup>25</sup> The converses of a) and b),  $PEp \rightarrow PWp$  and  $PE\sim p \rightarrow BWp$  are very implausible, so we will pass over them.

Abstractly, this says every option is either completely indifferent morally, or it has the strongest positive deontic and aretaic statuses and its negation has the strongest negative aretaic (not only deontic) status, or it has the strongest negative deontic and aretaic statuses and its negation has the strongest positive aretaic (not only deontic) status. Clearly with c)-f) conjoined, the aretaic and deontic categories have collapsed extensionally:

$$\begin{aligned} \text{OBp} &\leftrightarrow (\text{PWp} \ \& \ \text{BW}\sim\text{p}) \\ \text{OPp} &\leftrightarrow \text{AIp} \\ \text{IMp} &\leftrightarrow (\text{PW}\sim\text{p} \ \& \ \text{BWp}), \end{aligned}$$

We also get these aretaic-to-aretaic “collapses”:

$$\begin{aligned} \text{PWp} &\leftrightarrow \text{BW}\sim\text{p} \\ \text{BWp} &\leftrightarrow \text{PW}\sim\text{p}, \end{aligned}$$

saying that a proposition reflects favorably on an agent exactly when its negation reflects unfavorably on her. e) and f) are both quite contentious. Consider e). It would seem that something might be obligatory for me even though it is not praiseworthy, for example, that I do not kill my neighbors’ children, even though I have no motive to do so or anything to gain, or even perhaps I do have some horrible motive, but refrain out of fear only. Am I praiseworthy in either case? f) is a bit more plausible, but it is often thought that one can non-culpably do the wrong/impermissible thing, especially with derivative prohibitions where I do not realize that some general prohibition I fully understand is unwittingly instantiated by my current action in my current circumstances.

Before turning back to c) and d), let me note that in addition to bridging principles like a)-f), there is also the issue of embedding of aretaic and deontic operators in the scope of one another. Suppose you kindly volunteer to drive me to the airport, and I accept and plan accordingly. You will then be *obligated* to do so, but mightn’t you then also be praiseworthy for being so obligated? Likewise, suppose you promise to drive me, but shouldn’t have, since it conflicts with some prior but *now* lesser commitment that you now must break but had no right to thwart in the first place. Won’t you then be obligated to take me, but blameworthy for being so obligated? On the face of it, iteration tracks important pre-theoretic normative phenomena, and although it is beyond the scope of this paper to explore, doing so seems to be more than an idle “logicians’ exercise”.

Turning back now to c) and d), I think c) has little plausibility on reflection despite its explicit or tacit endorsement in some early literature tacitly or explicitly hostile to supererogation. Supererogatory acts are counterexamples, and note furthermore that I may have many mutually exclusive praiseworthy options, like saving Tiny Tim or Tiny Tara from the burning building, where it is not even possible to save both. I think d) is more attractive and still recently endorsed by some (e.g. [20] and [2]), but we have already tacitly rejected it in rejecting the prior bridging principle, b). For there I argued that I might be blameworthy for refusing to help a friend, even though my circumstances have changed unbeknownst to me in such a way that it is indeed not only permissible but obligatory for me to refuse to help the friend out (else my children may starve). Here refusing appears to be obligatory and so permissible, yet blameworthy.<sup>26</sup>

Still I think d) may be behind the widespread rejection of offenses. Although supererogation has been a marginalized concept in ethical theory and deontic logic, it is nonetheless much less controversial than that of offenses (suberogation). Even the staunchest defenders of supererogation (see [6] and [16]) raise serious doubts about offenses and argue against the alleged symmetry between offense and supererogation. One standard line of rejection is that if you allow for a full mirror image of supererogation, then not only can an option be blameworthy yet permissible, but it can be *blameworthy to the highest degree* and still be permissi-

---

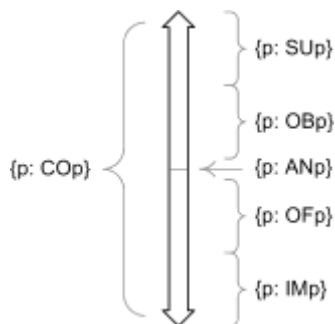
<sup>26</sup> Ish Haji and I provided further arguments against this principle (as well as c)), and its use in debates about the principle of alternative possibilities, responsibility, and determinism in “Supererogation and Frankfurt Examples”. Northern New England Philosophical Association presentation, 2009. See also [5] and [21].

ble, since a supererogatory option can be praiseworthy to the highest degree and still be permissible to skip. But then a downright reprehensible action could turn out to be permissible, which is deemed surely false. The objection is that once you open the door to allowing permissible blameworthy options like offenses, it seems hard to find any principled way to put any limit on the degree of blameworthiness that might be permissible in some circumstances. [16] offers two additions (QS and QO) to mitigate against the complaints of some anti-supererogationists who also invariably reject offenses, and he himself thinks there are better reasons to doubt offenses than to doubt supererogation, but he there overlooks the fact that the appeal of a very simple aretaic-deontic bridge principle like d) may be motivating the especially widespread rejection of offenses, one that would at once lead to a rejection of both of his quasi notions. In later work, he is clearer about this (e.g. [17]). However, our earlier reflections on b) suggest that a person can indeed be *extremely* blameworthy for an action that might not only be permissible but downright obligatory. Thus our reflections appear to suggest that this pathway for arguing against offenses (or Mellema's two quasi-notions) may very well rest on a simple though widely endorsed false presupposition.

Let me also add here that denying d) also allows us to make sense of a puzzle for deontic logic: "owed favors" [4]: we often say things to imply that we *owe* someone a *favor*, that a favor is due, etc. At a glance, it would seem that if it's a favor, it is not obligatory; but if it is owed, it is called for, and perhaps thus obligatory after all. What to do? Denying d) provides us with the option of saying that although a favor might not be *obligatory*, we might nonetheless be *blameworthy* for not doing a favor in some cases.<sup>27</sup>

[14] and [19] both endorsed theses that are natural errors when first reflecting on supererogation. The main one I will call "the Ranking Thesis" (RT):

$$(SU^a p \ \& \ OBq \ \& \ ANr \ \& \ OF^a s \ \& \ IMt) \rightarrow (p > q > r > s > t).$$



[1] makes clear the problems with this view, which problems we here summarize by noting that a small favor like lending a book you've already read to a friend might be supererogatory and nice but not warrant high praise, whereas sometimes our obligations are truly arduous to fulfill and very tempting to shirk (more below on this). Thus it will often be more praiseworthy to do something obligatory than to do something supererogatory. Similarly, if there are offenses, then there is no reason to think that these cannot sometimes be of such a caliber as to reflect much more poorly on an agent than shirking some trivial obligation.

The Ranking Thesis is a version of what is called the "continuity requirement" (see [6] and [16]): that supererogatory acts must involve the same value that makes actions obligatory (thus the continuity), but more

<sup>27</sup> An alternative purely deontic solution is possible through DWE, though not SDL: an "owed" favor is one that you *ought* to give, but is not one you *must* give (it is optimal but not required). For arguments that these are distinct and that DWE models them, see [8].

of that value. Although it is possible to give one interpretation of the orderings in DWE that at least allows one to make sense of this enough to warrant a serious discussion of some version of this thesis for *doing more than the minimum*, nonetheless, in the context of the aretaic ordering, it has little plausibility, since what makes an action praiseworthy (and not blameworthy to skip) has to do with where the action (inaction) lies on the scale of *aretaic* appraisal of the *agent*, whereas where the action lies on a scale that deontically evaluates the act as obligatory or as the minimum or whatever looks like a very different scale that evaluates actions not agents. Although we cannot explore this in depth here, our remarks in the last section will reinforce doubts about the continuity requirement for supererogation.

[1] also rejects Meinong's "laws of omission", which we can partially illustrate here by:

$$SU^a p \leftrightarrow OF^a \sim p.$$

If it is supererogatory for me to jump on a grenade to save my equally-able-to-jump comrades, it is not thereby blameworthy for me to not do so. Similarly, if it is an offense for me to not say "hello" back when passing you in a rush in the hall, it is not supererogatory for me to do so. Behind these laws of Meinong are probably more fundamental aretaic mistakes (e.g.  $PWp \leftrightarrow BW \sim p$ ) but we will study Schwartz', Meinong's, and Chisholm's schemes and the continuity requirement elsewhere.

In the next section we show that the classical analysis of the most plausible and widely accepted of the first four notions defined above, is fundamentally flawed.

## 7 A Basic Flaw in the Classical Analysis of Supererogation

As mentioned, my prior work in DWE does not involve any aretaic operators, and thus can't express supererogation, since that entails praiseworthiness, but I want to now suggest that this gap cuts both ways, and that DWE has the better of it: the classical conception of supererogation presupposes one of the deontic non-aretaic concepts of DWE, one central to DWE's success in accounting for an arguably more fundamental cousin of supererogation.

First, note that an action can be obligatory and highly praiseworthy. Consider a soldier on point. She stands her appointed ground faithfully in face of a sudden enemy attempt to overrun the main camp, despite extreme danger of death. Second, notice that if something is obligatory, then doing the least you can do involves doing that thing, and sometimes the minimum you can permissibly do is the same as what is obligatory--there are no graded options to speak of. It follows from the preceding case, along with the DWE principle that  $p$  is obligatory entails that  $p$  is involved in doing the minimum, that it can be praiseworthy to do the minimum. Now the crucial question: *Can it be praiseworthy to do the minimum even when one can also go beyond the call by doing more than the bare minimum?* Yes. Consider a minor variant of our prior case. Suppose there is now a first and a second position on point, the second being a slightly safer fallback position but also slightly riskier for the camp, so the first is better all in all. Now suppose it is permissible to pick either spot to make a stand for whoever is on point, by agreement of the group, etc. In face of a discernibly massive enemy attack, our soldier in good faith picks and holds fast the second position, at risk of near certain death. Here the least she can do is hold the second position. Obviously she "could" also retreat, hide, or play dead, but not permissibly so. Still the temptation to take the latter sort of impermissible option might be very very intense and such that many would do that. It can then surely be praiseworthy for her to hold even the second position in such circumstances. But now notice two other things about this case: 1) It is deontically *optional for her to hold the second position*: for she can also hold the first position instead, thereby going beyond the call. 2) It is also *not blameworthy for the agent to not hold the second position*, for then if she went beyond the call by holding the first position for the best of reasons, since that precludes holding the second, she would thereby be blameworthy as a consequence, which is not plausible.

So here we have an action that satisfies all three conditions of the classical analysis of supererogation, yet it is not beyond the call; indeed the action is *the minimum required*. When we can do more good than we

have to, doing the minimum will always be optional, as DWE ratifies, and it can't be automatically blameworthy to not do the minimum, for then going beyond the call would invariably entail something blameworthy, which is absurd. And sometimes our permissible choices are arduous enough that even taking the minimally permissible ones in good faith are praiseworthy. The soldier's holding the 2<sup>nd</sup> point is not intuitively supererogatory, and fails to fall under the classical *conception* of supererogation, so the classical *analysis*, despite its pervasiveness, is flawed: it does not give sufficient conditions for its target class of acts. Formally, we have argued that the following is satisfiable:

$$MIp \ \& \ PWp \ \& \ \sim BW\sim p \ \& \ OPp$$

and this, coupled with the classical analysis, then entails the satisfiability of

$$MIp \ \& \ SU^a p$$

but if we read the latter operator as intended, then the result is surely unpalatable and contrary to the classical conception that the classical analysis is intended to capture. Given the intended interpretation, we must reject the classical analysis:

$$SU^a p \stackrel{\text{def}}{=} PWp \ \& \ \sim BW\sim p \ \& \ OPp$$

In particular, the *deontic* condition is too weak for supererogation. Focusing only on actions that are aretically praiseworthy to do and not blameworthy to omit, and then merely adding deontic optionality is insufficient to guarantee they are of the intended kind. Put another way, such an act can be optional for the wrong reasons: because it is the least the agent can do and there are better options--it is a *surpassable minimum*!

There is some further irony in this. Supererogationists from Urmson forward accused deontic logic of ruling out supererogatory actions. I have argued that this is due to a conflation of deontic *optionality* with deontic *indifference* by both deontic logicians and ethical theorists, but our reflections today suggest that the classical analysis can't capture its intended class of actions without returning to deontic logic and *increasing* the deontic notions it relies on. Of course it needs more than the deontic concept of optionality that SDL can provide. It also needs, at the very least, the deontic concept of *doing the minimum that morality demands*. This concept is not expressible in SDL, but it is in DWE, which expands considerably upon the expressive resources of SDL, allowing for the expression of a person's exceeding the minimum that morality demands, and not entailing that this need be praiseworthy, for in fact it need not be. One can do more than the minimum for the wrong reasons or even for bad reasons and not be praiseworthy at all. For example, I rescue an infant from a burning building thinking the plans to bomb a school bus-load of children tomorrow are hidden in the infant's diapers, and I am solely motivated by the desire to get those plans and use them (but destined to fail, say). By saving the child, I do more good than I had to do, but I am surely not praiseworthy. I would suggest that the more objective and act-evaluative notion of doing more than the minimum is both more fundamental to our moral scheme and more important to it. This is reflected in the fact that it looks like the required revision of the classical analysis needed to capture the classical conception of supererogation is underscored in this revised analysis:

Revision of Classical Analysis of Supererogation:  $SUP \stackrel{\text{def}}{=} OPp \ \& \ PWp \ \& \ \sim BW\sim p \ \& \ MI\sim p.$

So something is genuinely supererogatory iff it meet the conditions of the classical analysis of supererogation and it is also precluded by doing the minimum. Given that in DWE,  $BCp \leftrightarrow (PEp \ \& \ MI\sim p)$  and  $MI\sim p \rightarrow PE\sim p$  are valid and derivable, it follows that  $BCp \leftrightarrow (MI\sim p \ \& \ OPp)$ , so we get:

SUp iff  $BCp \ \& \ PWp \ \& \ \sim BW\sim p$ .

We must thus of course also reject the regularly embraced equivalence:

The Equivalence: An act is Supererogatory iff it is *Beyond the Call* (of duty),

For being an action beyond the call is a necessary but not sufficient condition for supererogation.

Let me note that if "offense" in the classical literature is meant to be the mirror image of supererogation, as is often stated, and as the alternative term "suberogation" strongly suggests, then the corresponding account would be:

Revision of Classical Analysis of Suberogation: OFp iff  $BWp \ \& \ \sim PW\sim p \ \& \ OPp \ \& \ MA\sim p$ .

And given the symmetry in DWE between PS (Permissible Suboptimality) and BC, since  $PSp \leftrightarrow (PEp \ \& \ MA\sim p)$  and  $MA\sim p \rightarrow PE\sim p$  are valid and derivable, it follows that  $PSp \leftrightarrow (MA\sim p \ \& \ OPp)$ , and so we get:

OFp iff  $PSp \ \& \ BWp \ \& \ \sim PW\sim p$ .

We leave for another occasion the controversial issue of whether the pre-theoretic conception of an offense is truly symmetrical with that of supererogation, merely noting here that if we are looking for a symmetrical analogue, the one above appears to be it. To be sure, it will not do to characterize an offense as something we ought not do ideally, but are permitted to do. That is, this equivalence must be rejected:

The Analogue Equivalence: An act is an offense/suberogatory iff it is permissibly suboptimal.

For in a variant of our earlier mailwoman woman rescue case, she might heroically save one child by entering the left side of the building, at the cost of not saving twins on the right side of the building at even great risk. Any rescue might be intuitively supererogatory. The former will be both supererogatory and permissibly suboptimal nonetheless--as argued elsewhere [8, 10], this is just a consequence of the fact that there can be supererogatory options available that are mutually exclusive and nonetheless ranked one above the other. But surely the rescue of even the one child is not an offense or suberogatory. Still, if an offense is intended to be the mirror image of supererogation, it had better always be a permissibly suboptimal act—one we can but ought not do—this is a necessary, even if not sufficient, condition.

All this suggests that in addition to exploring FAA (and kin) in various directions, linking FAA with DWE more thoroughly is also in order. The result will be a considerably enriched system, and one which will distinguish between two closely related and often conflated concepts, supererogation and doing more good than one has to do, and between a permissibly suboptimal act and an offense (or suberogatory act), as well as allowing many other potential distinctions with the blending of such aretaic and deontic resources, and the combinatorial boost in expressive power this yields.<sup>28</sup>

---

<sup>28</sup> I benefited from comments from four anonymous referees on a shorter predecessor submitted for DEON'08, and from three referees for JAL on the penultimate version of the present article. I also benefited from Marek Sergot, Jorg Hanson, Andrew Jones, and Peter Varanas, among others, in discussions at DEON'08, from Lou Goble and Peter Scotch in discussions at the Society for Exact Philosophy 2009 meeting in Edmunton, and from Ishtiyaque Haji in correspondence about related issues.

## Appendix: Formal Semantics for DWE

The following formal framework is generalized in [7], where the metatheorem below is proven as a special case.

Let  $F = \langle W, A, \leq \rangle$ , be a DWE-Frame, where:

- (1)  $W \neq \emptyset$
- (2)  $A \subseteq W^2$  and  $\forall i \exists j Aij$  (seriality)
- (3)  $\leq \subseteq W^3$ : a)  $(k \leq_i j \text{ or } j \leq_i k) \text{ iff } (Aij \ \& \ Aik)$ , for any  $i, j, k$  in  $W$   
b) if  $j \leq_i k$  and  $k \leq_i l$  then  $j \leq_i l$ , for any  $i, j, k, l$  in  $W$ .

$P$  is an Assignment on  $F$ :  $F = \langle W, A, \leq \rangle$  is a DWE-Frame and  $P$  is a function from the propositional variables (PV) to  $\text{Power}(W)$ , defined on PV.

$M = \langle F, P \rangle$  is a DWE-Model:  $F = \langle W, A, \leq \rangle$  is a DWE-frame and  $P$  is an assignment on  $F$ .

Truth at a World in a DWE-Model: Let  $M = \langle F, P \rangle$  be a DWE-model, where  $F = \langle W, A, \leq \rangle$  and  $j =_i k \text{ =df } j \leq_i k \ \& \ k \leq_i j$ . Then for any  $i$  in  $W$ :

Basic Truth-Conditions at a world  $i$ , in a model,  $M$ :

- 0) (Usual Conditions for variables and truth functional connectives)
- 1)  $M \models_i \mathbf{OB}p$ :  $\forall j(\text{if } Aij \text{ then } M \models_j p)$ .
- 2)  $M \models_i \mathbf{MA}p$ :  $\exists j(Aij \ \& \ (k)(\text{if } j \leq_i k \text{ then } M \models_k p))$ .
- 3)  $M \models_i \mathbf{MI}p$ :  $\exists j(Aij \ \& \ (k)(\text{if } k \leq_i j \text{ then } M \models_k p))$ .
- 4)  $M \models_i \mathbf{IN}p$ :  $\forall j[\text{if } Aij \text{ then } \exists k(k =_i j \ \& \ M \models_k p) \ \& \ \exists k(k =_i j \ \& \ M \models_k \sim p)]$ .

Derivative Truth Conditions:

- 5)  $M \models_i \mathbf{PE}p$ :  $\exists j(Aij \ \& \ M \models_j p)$ .
- 6)  $M \models_i \mathbf{IM}p$ :  $\forall j(\text{if } Aij \text{ then } M \models_j \sim p)$ .
- 7)  $M \models_i \mathbf{OM}p$ :  $\exists j(Aij \ \& \ M \models_j \sim p)$ .
- 8)  $M \models_i \mathbf{OP}p$ :  $\exists j(Aij \ \& \ M \models_j p) \ \& \ \exists j(Aij \ \& \ M \models_j \sim p)$ .
- 9)  $M \models_i \mathbf{SI}p$ :  $\exists j[Aij \ \& \ \text{either } \forall k(\text{if } k =_i j \text{ then } M \models_k p) \ \text{or } \forall k(\text{if } k =_i j \text{ then } M \models_k \sim p)]$ .
- 10)  $M \models_i \mathbf{BC}p$ :  $\exists j(Aij \ \& \ M \models_j p) \ \& \ \exists j[Aij \ \& \ \forall k(\text{if } k \leq_i j \text{ then } M \models_k \sim p)]$ .
- 11)  $M \models_i \mathbf{PS}p$ :  $\exists j(Aij \ \& \ M \models_j p) \ \& \ \exists j[Aij \ \& \ \forall k(\text{if } j \leq_i k \text{ then } M \models_k \sim p)]$ .

Truth in a DWE-Model:  $M \models p$  iff  $M \models_i p$ , for every  $i$  in  $W$  of  $M$ .

Validity for a Set of DWE-Models:  $C \models p$  iff  $M \models p$ , for all  $M$  in  $C$ .

Metatheorem: The DWE-logic is determined by the class of DWE-models.

## References

1. Chisholm, R.M., Supererogation and Offence: A Conceptual Scheme for Ethics. *Ratio* 5, 1963: 1-14.
2. Copp, D. Defending the Principle of Alternate Possibilities: Blameworthiness and Moral Responsibility. *Nous* 31, 1997: 441-456.
3. Demolombe, Robert and Andrew J. I. Jones. Actions and normative positions: a modal-logical approach. *A Companion to Philosophical Logic*, Dale Jacquette (ed.). Blackwell. 2002: 694-705.
4. Driver, J. The Suberogatory. *Journal of Philosophy* 70(3), 1992: 286-295.
5. Haji, I. *Moral Appraisability: Puzzles, Proposals, and Perplexities*. New York: Oxford University Press, 1998.
6. Heyd, D., *Supererogation: Its Status in Ethical Theory*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press, 1982.
7. Mares, E.D. and P. McNamara. Supererogation in Deontic Logic: Metatheory for DWE and Some Close Neighbours. *Studia Logica* 59(3), 1997: 397-415.
8. McNamara, P. Doing Well Enough: Toward a Logic for Commonsense Morality. *Studia Logica* 57(1), 1996: 167-192.
9. McNamara, P. Must I Do What I Ought? (or Will the Least I Can Do Do?). *Deontic Logic, Agency and Normative Systems*, Brown M. and J. Carmo (eds.) 1996, Springer Verlag: New York: 154-173.
10. McNamara, P. Making Room for Going Beyond the Call. *Mind* 105(419), 1996: 415-450.
11. McNamara, P. Doing Well Enough in an Andersonian-Kangerian Framework. *Norms, Logics and Information Systems: New Studies in Deontic Logic and Computer Science*, P. McNamara and H. Prakken (eds.) 1999, IOS Press: Washington, DC: 181-198.
12. McNamara, P. Toward a Framework for Agency, Inevitability, Praise and Blame. *Nordic Journal of Philosophical Logic* 5(2), 2000: 135-160.
13. McNamara, P. Agential Obligation as Non-Agential Personal Obligation plus Agency. *Journal of Applied Logic* 2(1), 2004: 117-152.
14. Meinong, A. *Psychologisch-ethische Untersuchungen zur Werth-Theorie*. Graz, 1894.
15. Mellema, G., Quasi-Supererogation. *Philosophical Studies* 52, 1987: 141-150.
16. Mellema, G., *Beyond the Call of Duty: Supererogation, Obligation, and Offence*. 1991, Albany: SUNY Press.
17. Mellema, G., Supererogation, Blame, and the Limits of Obligation. *Philosophia* 24(1-2), 1994: 171-182.
18. Pybus, E.M., Saints and Heroes. *Philosophy* 57, 1982: 193-200.
19. Schwarz, E. *Über den Wert, das Soll, und das richtige Werthalten*. Graz, 1934.
20. Widerker, D., Frankfurt on 'Ought Implies Can' and Alternative Possibilities. *Analysis* 51, 1991: 222-224.
21. Zimmerman, M. J. A Plea for Accuses. *American Philosophical Quarterly* 34, 1997: 229-43.