



# The irreducibility of collective obligations

Allard Tamminga<sup>1,2</sup> · Frank Hindriks<sup>1</sup>

Published online: 12 January 2019  
© The Author(s) 2019, corrected publication 2019

**Abstract** Individualists claim that collective obligations are reducible to the individual obligations of the collective’s members. Collectivists deny this. We set out to discover who is right by way of a deontic logic of collective action that models collective actions, abilities, obligations, and their interrelations. On the basis of our formal analysis, we argue that when assessing the obligations of an individual agent, we need to distinguish individual obligations from member obligations. If a collective has a collective obligation to bring about a particular state of affairs, then it might be that no individual in the collective has an individual obligation to bring about that state of affairs. What follows from a collective obligation is that each member of the collective has a member obligation to help ensure that the collective fulfills its collective obligation. In conclusion, we argue that our formal analysis supports collectivism.

**Keywords** Collective obligation · Collective responsibility · Collectivism · Individualism · Deontic logic · Game theory

---

✉ Allard Tamminga  
a.m.tamminga@rug.nl  
Frank Hindriks  
f.a.hindriks@rug.nl

<sup>1</sup> Faculty of Philosophy, University of Groningen, Oude Boteringestraat 52, 9712 GL Groningen, The Netherlands

<sup>2</sup> Department of Philosophy and Religious Studies, Utrecht University, Janskerkhof 13, 3512 BL Utrecht, The Netherlands

## 1 Introduction

Most benefits and harms in the socio-economic realm are brought about by collectives rather than individuals. Individuals can be held morally responsible for their actions, but does it also make sense to hold a collective morally responsible for the things it does or fails to do? An affirmative answer immediately raises the question of whether collective moral responsibility is reducible to the individual moral responsibility of the collective's members. Individualists say yes, collectivists say no. If the individualists are right, what kind of reduction proves them right?

Two kinds of reduction are relevant here: (1) *logical reduction* concerns whether statements about collectives are logically equivalent to conjunctions of statements about individuals, and (2) *ontological reduction* concerns whether collective entities are just sets of individuals. For argumentative purposes, we start from the assumption that ontological reductionism is true, that is, we assume that a group is a set of individuals and that a group action is an ordered set of individual actions.<sup>1</sup> We argue that, even if ontological reductionism is true, logical reductionism is false.

In this paper, we focus on the notion of collective obligation. Obligations are forward-looking responsibilities. They are also a good place to start for studying the reducibility of backward-looking responsibilities, at least in so far as blameworthiness is concerned. Not fulfilling an obligation is a necessary condition for being individually blameworthy: if an individual agent is individually blameworthy, then she failed to fulfill an obligation.<sup>2</sup> Similarly, not fulfilling a collective obligation is a necessary condition for being collectively blameworthy: if a group is collectively blameworthy, then it failed to fulfill a collective obligation. Because of this, we employ a two-step strategy to establish whether collective moral responsibility is logically reducible to individual moral responsibility. We first develop a formal theory of collective obligations and then use the formal properties we have identified to find out how to evaluate the actions of individual group members in case a group of agents does not fulfill its collective obligation.

The structure of our paper is as follows. In Sect. 2, we explain why, in studying collective moral responsibility, it is important to develop a logical theory of collective actions and obligations. In Sect. 3, we present our deontic logic of collective action. This deontic logic is then used to study (1) logical relations between collective actions and individual actions, and (2) logical relations between collective obligations and individual obligations. We list some intuitive principles concerning collective actions and obligations and use our deontic logic to determine whether or not they are valid. In Sect. 4, we submit that when assessing the obligations of an individual agent, we must distinguish *individual* obligations from *member* obligations. We argue that, although collective obligations are *not* logically

---

<sup>1</sup> Copp (2006), List and Pettit (2011), and Hindriks (2013) argue against ontological reductionism about structured groups.

<sup>2</sup> Not fulfilling an obligation, however, is not a sufficient condition for being individually blameworthy: the agent might have a justification or an excuse for not doing her duty. Furthermore, whereas blame requires not fulfilling an obligation, supererogatory actions make it clear that praise does not require fulfilling an obligation.

reducible to individual obligations, given some well-defined assumptions they do imply member obligations. What follows from a collective obligation is that each individual has a member obligation to help ensure that the collective fulfills its collective obligation. In conclusion, we argue that, because member obligations presuppose collective obligations, our formal analysis supports collectivism.

## 2 Logical puzzles

The study of collective moral responsibility would be facilitated by a formal theory that systematically distinguishes valid principles about collective actions and obligations from invalid ones. Unfortunately, a formal theory of this sort has yet to be offered. Most of the existing literature on collective moral responsibility remains informal and relies on everyday examples. Feinberg (1968) discusses the Jesse James train robbery and concludes that the passengers are collectively responsible for not having prevented it. Held (1970) takes us to a subway car in which five bystanders watch someone being strangled and also concludes that the collective is responsible.<sup>3</sup> Sometimes such everyday examples are used as counterexamples to show that specific logical principles connecting collective actions and obligations with individual ones are *invalid*.<sup>4</sup> So far so good.

But are there also *valid* principles that connect collective actions and obligations with individual ones? Frank Jackson thinks there are: ‘it is surely unbelievable that there are *no* valid principles linking the moral status of group acts with the moral status of constituent individual acts’ (Jackson 1987, p. 107). If we try to show that Jackson is right, we are presented with a difficulty: although everyday examples can be used to argue that some logical principles are *invalid*, they cannot be used to argue that some such principles are *valid*. To identify such validities, and also to systematize the scattered invalidities that were identified with the help of everyday examples, we need a formal theory of collective actions and obligations. This is by no means a straightforward matter.

Various authors appear to be puzzled about such a logic of collective actions and obligations. Let us consider two examples. (1) Kutz (2000a, p. 105) asserts that it is ‘puzzling’ that the conditional ‘If agent *i* and agent *j* together see to it that  $\phi$ , then agent *i* sees to it that  $\phi$  and agent *j* sees to it that  $\phi$ ’ is invalid, ‘while ordinarily each conjunct of a true conjunction is also true’. (2) Lawford-Smith (2012, pp. 455–456)

<sup>3</sup> Feinberg argues that such collective responsibility cannot be reduced to individual responsibility, whereas Held defends reductionism about collective responsibility.

<sup>4</sup> Held argues that ‘from the judgment “Collectivity *C* ought (ought not) to have done *A*”, judgments of the form “Member *M* of *C* ought (ought not) to have done *A*” cannot be derived’ and that ‘judgments about the moral responsibility of its members are not logically derivable from judgments about the moral responsibility of a collectivity’ (Held 1970, pp. 474–475). Compare Gilbert: ‘What does the blameworthiness of the collective’s act imply about the personal blameworthiness of any one member of that collective? From a logical point of view, the short answer is: *nothing*. Everything depends on the details of a given member’s particular situation’ (Gilbert 2006, p. 109). Likewise, Isaacs argues that ‘claims about collective moral responsibility neither entail nor are derivable from claims about individual moral responsibility’ (Isaacs 2011, p. 24).

maintains that the conditional ‘If agent  $i$  ought to see to it that  $\phi \wedge \psi$ , then agent  $i$  ought to see to it that  $\phi$  and agent  $i$  ought to see to it that  $\psi$ ’ and the conditional ‘If agent  $i$  and agent  $j$  together ought to see to it that  $\phi$ , then agent  $i$  ought to see to it that  $\phi$  and agent  $j$  ought to see to it that  $\phi$ ’ are both substitution instances of the logical principle of the distribution of ‘oughts’ over a conjunction. Because she proposes that we reject the second conditional, she questions the principle’s validity. The deontic logic of collective action that we present in Sect. 3 does of course vindicate the validity of the logical principle that the truth of a conjunction implies the truth of both of its conjuncts. Nonetheless, our deontic logic indicates why Kutz’s conditional is invalid, why Lawford-Smith’s first conditional is valid, and why her second conditional is not. That we can draw such conclusions illustrates the added value of our logic.

In Sects. 3 and 4 we develop a formal framework to study the logical relations between collective actions and obligations on the one hand and individual actions and obligations on the other. Our framework will remove puzzlement about the formal properties of collective action and obligation, and, more generally, provide a tool to investigate the formal relations between collective and individual responsibility. The framework will therefore also help us to achieve our second goal, which is to clarify the concept of logical reduction in the debate on individualism versus collectivism about collective moral responsibility. For these reasons, we construct a deontic logic of collective action.

### 3 A deontic logic of collective action

The deontic logic of collective action presented in this section is inspired by John Horty’s (1996, 2001) deontic logic of agency.<sup>5</sup> The present section and the next extend Horty’s ideas to a formal theory of collective obligations. Our theory provides clear answers to central questions about the relation between collective and individual obligations. Can it be the case that an individual in a collective ought to bring about a certain state of affairs without its being the case that that collective itself ought to bring about that state of affairs? (Yes.) Can it be the case that a collective ought to see to it that a certain state of affairs obtains and that an individual in that collective ought to see to it that that state of affairs does not obtain? (No.) Are the obligations of a group and all of its subgroups consistent? (Yes.)

To investigate formal relations between collective and individual obligations, we make two central idealizing assumptions to filter out factors that interfere with these relations to the point of trivializing them. Because there are no necessary relations between obligations generated by different moral codes or between obligations at different moments in time, we study, first, collective and individual obligations that

---

<sup>5</sup> Our deontic logic of collective action stands in the tradition of the *stit* (‘sees to it that’) logics of agency developed by Kanger (1957), Pörn (1970), Kanger (1972), Hansson (1986), von Kutschera (1986), Horty and Belnap (1995), Belnap et al. (2001), Horty (2001), Kooi and Tamminga (2008), Carmo (2010), and Tamminga (2013).

are generated by the same moral code and, secondly, collective and individual obligations that hold at a single moment in time.

We use the term ‘group’ for any random collection of individuals (Held 1970). As a consequence, any combination of individual actions performed by such a random collection of individuals will be considered a group action performed by that random collective. Naturally, relaxing these assumptions may well result in different answers to the main questions about the relation between collective and individual obligations. It should therefore be borne in mind that our deontic logic of collective action is only a first step towards a fully-fledged formal analysis of collective actions and obligations.

Our deontic logic is, however, meant to be general in the following respect. Some group actions are just sets of individual actions. Think, for instance, of harming the ozone layer. Other group actions involve collective intentions, such as the collective intention to move the piano upstairs together. To attain a level of generality that enables us to study common features of group actions that involve collective intentions (like moving the piano upstairs) and of group actions that lack them (like harming the ozone layer), we initially assume that the sets of individuals who perform group actions have not formed a collective intention. On the basis of this assumption, we model group actions as sets of individual actions. Later on, we relax this assumption to accommodate joint actions generated by collective intentions as well.<sup>6</sup> We model collective intentions in a minimal way as group plans that have been adopted by all members of the group.

This technical section proceeds as follows. First, we define our formal language and, by way of illustration, translate the statements that puzzled Kutz and Lawford-Smith into this language. Secondly, we define the deontic game models that interpret the formal language and then check whether the translations of Kutz’s and Lawford-Smith’s statements are valid, that is, whether the translations of these statements are true in all models. Thirdly, we justify the answers to the questions that opened this section by checking the validity of some general claims about relations between group actions and subgroup actions and between group obligations and subgroup obligations. All claims are proved in the “Appendix”.

### 3.1 Formal language

Our propositional modal language  $\mathcal{L}$  is built from a finite set  $N = \{i_1, \dots, i_n\}$  of individual agents and a countable set  $\mathfrak{P} = \{p_1, p_2, \dots\}$  of atomic formulas. The formal language  $\mathcal{L}$  is the smallest set (in terms of set-theoretical inclusion) that satisfies conditions (i) through (iv):

- (i)  $\mathfrak{P} \subseteq \mathcal{L}$
- (ii) If  $\phi \in \mathcal{L}$ , then  $\neg\phi \in \mathcal{L}$  and  $\diamond\phi \in \mathcal{L}$
- (iii) If  $\phi \in \mathcal{L}$  and  $\psi \in \mathcal{L}$ , then  $(\phi \wedge \psi) \in \mathcal{L}$

<sup>6</sup> Collective intentions play a central role in the analysis of joint actions. See, for instance, Tuomela and Miller (1988), Gilbert (1990), Searle (1990), and Bratman (1992). For analyses of collective action that lack collective intentions, see Jackson (1987), Kutz (2000b), and Chant (2007).

- (iv) If  $\mathcal{G} \subseteq N$  and  $\phi \in \mathfrak{L}$ , then  $[\mathcal{G}]\phi \in \mathfrak{L}$  and  $(\mathcal{G})\phi \in \mathfrak{L}$ .

We omit parentheses, brackets, and braces if the omission does not give rise to ambiguities. The operators  $\vee$ ,  $\rightarrow$ , and  $\Box$  abbreviate the usual constructions.

This propositional modal language can be used to express *alethic statements* like ‘It is possible that  $\phi$ ’ (formalized as  $\Diamond\phi$ ), *agentive statements* like ‘Group  $\mathcal{G}$  of agents sees to it that  $\phi$ ’ (formalized as  $[\mathcal{G}]\phi$ ), and *deontic statements* like ‘Group  $\mathcal{G}$  of agents ought to see to it that  $\phi$ ’ (formalized as  $(\mathcal{G})\phi$ ). Abilities can then be expressed by a combination of operators: ‘Group  $\mathcal{G}$  of agents has the ability to see to it that  $\phi$ ’ is formalized as  $\Diamond[\mathcal{G}]\phi$ .<sup>7</sup>

The statements that puzzled Kutz (2000a, p. 105) and Lawford-Smith (2012, pp. 455–456) are:

- (1) If agent  $i$  and agent  $j$  together see to it that  $\phi$ , then agent  $i$  sees to it that  $\phi$  and agent  $j$  sees to it that  $\phi$ .
- (2) If agent  $i$  ought to see to it that both  $\phi$  and  $\psi$ , then agent  $i$  ought to see to it that  $\phi$  and agent  $i$  ought to see to it that  $\psi$ .
- (3) If agent  $i$  and agent  $j$  together ought to see to it that  $\phi$ , then agent  $i$  ought to see to it that  $\phi$  and agent  $j$  ought to see to it that  $\phi$ .

They can now be translated as (1’), (2’), and (3’):

- (1’)  $[i, j]\phi \rightarrow ([i]\phi \wedge [j]\phi)$
- (2’)  $(i)(\phi \wedge \psi) \rightarrow ((i)\phi \wedge (i)\psi)$
- (3’)  $(i, j)\phi \rightarrow ((i)\phi \wedge (j)\phi)$ .

To show that (2’) is valid and that (1’) and (3’) are not, we need to know the conditions under which a formula of our modal language is true. These truth-conditions are specified by our deontic game models.

### 3.2 Deontic game models

What ingredients do we need to specify truth-conditions for agentive and deontic statements? Intuitively, the truth of an agentive statement like ‘Group  $\mathcal{G}$  of agents sees to it that  $\phi$ ’ requires that group  $\mathcal{G}$  choose and perform a group action from among a range of group actions that are available to the group and that by performing this group action the group ensures that  $\phi$  is true. Moreover, if we order the available group actions in terms of how well they promote a given moral code, then, intuitively, the truth of a deontic statement like ‘Group  $\mathcal{G}$  of agents ought to see to it that  $\phi$ ’ requires that every best group action that is available to group  $\mathcal{G}$  ensures, if performed, that  $\phi$  is true.

To convert these intuitive requirements for the truth of agentive and deontic statements into truth-conditions, we use a model theory based on the possible-worlds semantics for *standard deontic logic* (Hilpinen 1971, pp. 13–15). A possible-worlds model for standard deontic logic involves a non-empty set of possible worlds, a set of possible worlds that are deontically ideal with respect to a given moral code (this set of deontically ideal possible worlds is a subset of the set

<sup>7</sup> For a formal analysis of the concept of individual ability, see Horty (2001, pp. 19–29).

of possible worlds), and a valuation function that assigns to each atomic formula a set of possible worlds where that formula is true. In such a possible-worlds model, a deontic statement like ‘It ought to be that  $\phi$ ’ is true if and only if  $\phi$  is true in all deontically ideal possible worlds.

Our *deontic game models* are possible-worlds models for standard deontic logic where the role of possible worlds is played by possible *action profiles* (to be defined below). Hence, a deontic game model consists of a set of possible action profiles, a set of possible action profiles that are deontically ideal with respect to a given moral code (this set of deontically ideal possible action profiles is a subset of the set of possible action profiles), and a valuation function that assigns to each atomic formula a set of possible action profiles where that formula is true.

Formally, a deontic game model involves a finite set  $N = \{i_1, \dots, i_n\}$  of individual agents. Each agent  $i$  in  $N$  is assigned a non-empty and finite set  $A_i$  of available individual actions.<sup>8</sup> We use  $a_i$  and  $a'_i$  as variables for actions in the set  $A_i$  of actions that are available to agent  $i$ . The set  $A$  of possible action profiles is given by the Cartesian product  $\times_{i \in N} A_i$  of all the individual agents’ sets of actions. Note that  $A (= \times_{i \in N} A_i)$  is non-empty, because all the  $A_i$ ’s are non-empty. We use  $a$  and  $a'$  as variables for action profiles in the set  $A$  of possible action profiles.<sup>9</sup> The set of deontically ideal possible action profiles (which is a subset of the set  $A$  of possible action profiles) is defined by a deontic ideality function  $d$  that assigns to each possible action profile  $a$  in  $A$  a value  $d(a)$  that is either 1 (if  $a$  is deontically ideal) or 0 (if  $a$  is not deontically ideal).<sup>10</sup> A valuation function  $v$  assigns to each atomic formula  $p$  in  $\mathfrak{P}$  a (possibly empty) set  $v(p)$  of possible action profiles.

As an illustration, let us define a deontic game model  $M_1$ , in which there are two agents  $i$  and  $j$ . Let  $A_i = \{a_i, a'_i\}$  and  $A_j = \{a_j, a'_j\}$ , that is, the actions that are available to agent  $i$  are  $a_i$  and  $a'_i$ , and the actions that are available to agent  $j$  are  $a_j$  and  $a'_j$ . The full set  $A$  of possible action profiles is then given by  $A_i \times A_j = \{(a_i, a_j), (a_i, a'_j), (a'_i, a_j), (a'_i, a'_j)\}$ . We define the deontic ideality function  $d$  as follows:

$$\begin{aligned} d(a_i, a_j) &= 1 \\ d(a_i, a'_j) &= 0 \\ d(a'_i, a_j) &= 0 \\ d(a'_i, a'_j) &= 1. \end{aligned}$$

Accordingly, the action profiles  $(a_i, a_j)$  and  $(a'_i, a'_j)$  are deontically ideal, and the action profiles  $(a_i, a'_j)$  and  $(a'_i, a_j)$  are not. Lastly, we define the valuation function  $v$  (the valuations of the other atomic formulas are left unspecified):

<sup>8</sup> We require finiteness for expository reasons only. If we were to admit infinite sets of individual actions, we would have to generalize the truth-condition for deontic statements accordingly—see footnote 12.

<sup>9</sup> We adopt the notational conventions of Osborne and Rubinstein (1994, Section 1.7).

<sup>10</sup> The deontic ideality function represents a given moral code. Our binary ordering of the action profiles in terms of deontic ideality can also be seen to reflect a simple preference ordering of agents who classify action profiles unambiguously as ‘good’ or ‘bad’. Deontic game models thus resemble Schelling’s (1960, p. 84) *pure-collaboration games* and Bacharach’s (2006, p. 122) *coordination contexts*.

$$v(p) = \{(a_i, a_j)\}$$

$$v(q) = \{(a_i, a_j), (a_i, a'_i), (a'_i, a'_j)\}.$$

Accordingly, the atomic formula  $p$  is true at action profile  $(a_i, a_j)$  and false at the other action profiles, and the atomic formula  $q$  is false at action profile  $(a'_i, a_j)$  and true at the other action profiles.

The deontic game model  $M_1$  can be pictured as in Fig. 1. The rows represent the actions that are available to agent  $i$ , and the columns represent the actions that are available to agent  $j$ . Each cell therefore represents an action profile. Moreover, each cell contains a number that indicates whether or not the corresponding action profile is deontically ideal. Finally, we use the convention that a cell contains an atomic formula if and only if that atomic formula is true at the corresponding action profile.

We now turn to group actions. Following the literature on the logic of collective action,<sup>11</sup> we think of an action performed by a group as a combination of individual actions performed by the group's members. In the present setting, we make the idealizing assumption that, conversely, any combination of individual actions performed by the members of any group is a group action performed by that group (we relax this latter assumption in Sect. 4, where we introduce plan-based collective intentions which allow us to make a distinction between aggregate behaviour and group agency).

Accordingly, for each non-empty group  $\mathcal{G} \subseteq N$  we define the set  $A_{\mathcal{G}}$  of group actions available to group  $\mathcal{G}$  as  $A_{\mathcal{G}} = \times_{i \in \mathcal{G}} A_i$ . We use  $a_{\mathcal{G}}$  and  $a'_{\mathcal{G}}$  as variables for group actions in the set  $A_{\mathcal{G}}$  of group actions that are available to group  $\mathcal{G}$ . Given an action profile  $a$  in a deontic game model and a group  $\mathcal{G} \subseteq N$ , we also use  $a_{\mathcal{G}}$  to refer to the combination of individual actions of  $\mathcal{G}$ 's members in  $a$ . Likewise, we use  $a_{-\mathcal{G}}$  to refer to the combination of individual actions of all of  $\mathcal{G}$ 's non-members in  $a$ . For example, given an action profile  $(a_i, a_j, a_k, a_l)$  in a four-agent deontic game model and the group  $\mathcal{G} = \{i, k\}$ , we have  $a_{\mathcal{G}} = (a_i, a_k)$  and  $a_{-\mathcal{G}} = (a_j, a_l)$ . In the same fashion, we use  $A_{-\mathcal{G}}$  to refer to  $\times_{i \in (N-\mathcal{G})} A_i$ .

Action profiles are fully determined by combinations of actions of all the individual agents. Hence, an agent only partly determines an action profile by performing an action, and this holds irrespective of whether the agent is an individual or a group. By performing an action an agent only makes a contribution to realizing an action profile. Nonetheless, by performing an action an agent can make statements true or false. In the present setting, given a specific action profile, a group sees to it that a particular statement is true if and only if the group's contribution to this action profile suffices to ensure the truth of this statement, regardless of the actions of the group's non-members. To be precise, given an action profile  $a$ , a group  $\mathcal{G}$  sees to it that  $\phi$  if and only if for all action profiles  $a'$  with  $a'_{\mathcal{G}} = a_{\mathcal{G}}$  it holds that  $\phi$  is true in  $a'$ . (The truth-condition for individual agency is obtained by filling in  $\{i\}$  for  $\mathcal{G}$ .)

<sup>11</sup> See Hansson (1986), Tuomela (1989), Belnap et al. (2001, Ch. 10), Horty (2001, pp. 30–33), and Carmo (2010).



	$a_j$	$a'_j$
$a_i$	$1/p, q$	$0/q$
$a'_i$	$0$	$1/q$

Fig. 1 Deontic game model  $M_1$

We model group obligations as follows: a group ought to see to it that a particular statement is true if and only if all of its optimal actions ensure the truth of that statement. To discover which of the actions available to a group are optimal, we order the available actions in terms of the deontic ideality of their possible action profiles. We thus transform the given binary deontic ideality ordering of the action profiles in  $A$  into an ordering of the actions available to group  $\mathcal{G}$ , that is, into an ordering of  $A_{\mathcal{G}}$ . The actions that are best in this ordering of  $A_{\mathcal{G}}$  are group  $\mathcal{G}$ 's optimal actions. Following Horty (2001, p. 68), we order a group's available actions in terms of *dominance*. Given two actions  $a_{\mathcal{G}}$  and  $a'_{\mathcal{G}}$  that are available to a group  $\mathcal{G}$ ,  $a_{\mathcal{G}}$  weakly dominates  $a'_{\mathcal{G}}$  (notation:  $a_{\mathcal{G}} \succeq a'_{\mathcal{G}}$ ) if and only if  $a_{\mathcal{G}}$  promotes deontic ideality at least as well as  $a'_{\mathcal{G}}$ , regardless of what the group's non-members do:

**Definition 1** (*Dominance*) Let  $M = \langle N, (A_i), d, v \rangle$  be a deontic game model. Let  $\mathcal{G} \subseteq N$  be a group of agents. Let  $a_{\mathcal{G}}, a'_{\mathcal{G}} \in A_{\mathcal{G}}$  be actions available to group  $\mathcal{G}$ . Then

$$a_{\mathcal{G}} \succeq a'_{\mathcal{G}} \text{ iff for all } a''_{-\mathcal{G}} \in A_{-\mathcal{G}} \text{ it holds that } d(a_{\mathcal{G}}, a''_{-\mathcal{G}}) \geq d(a'_{\mathcal{G}}, a''_{-\mathcal{G}}).$$

Strong dominance is defined in terms of weak dominance:  $a_{\mathcal{G}} \succ a'_{\mathcal{G}}$  if and only if  $a_{\mathcal{G}} \succeq a'_{\mathcal{G}}$  and  $a'_{\mathcal{G}} \not\succeq a_{\mathcal{G}}$ .

A group's optimal actions are the ones that are not strongly dominated:

**Definition 2** (*Optimality*) Let  $M = \langle N, (A_i), d, v \rangle$  be a deontic game model. Let  $\mathcal{G} \subseteq N$  be a group of agents. Then the set of  $\mathcal{G}$ 's *optimal actions*, denoted by  $Optimal(\mathcal{G})$ , is given by

$$Optimal(\mathcal{G}) = \{a_{\mathcal{G}} \in A_{\mathcal{G}} : \text{there is no } a'_{\mathcal{G}} \in A_{\mathcal{G}} \text{ such that } a'_{\mathcal{G}} \succ a_{\mathcal{G}}\}.$$

Finally, we need to specify truth-conditions for the formulas of our propositional modal language. We write  $M, a \models \phi$  if a formula  $\phi$  is true at an action profile  $a$  of a deontic game model  $M$ . The truth-conditions are given inductively:<sup>12</sup>

**Definition 3** (*Truth-conditions*) Let  $M = \langle N, (A_i), d, v \rangle$  be a deontic game model. Let  $\mathcal{G} \subseteq N$  be a group of agents. Let  $a \in A$  be an action profile. Let  $p \in \mathfrak{P}$  be an atomic formula and  $\phi, \psi \in \mathfrak{L}$  arbitrary formulas. Then

<sup>12</sup> If we were to admit infinite sets of individual actions (and hence admit infinite sets of action profiles), then the truth-condition for deontic statements would be as follows:  $M, a \models (\mathcal{G})\phi$  if and only if for all  $a' \in A$  with  $M, a' \not\models \phi$  there is an  $a'' \in A$  with  $M, a'' \models \phi$  such that (i)  $a''_{\mathcal{G}} \succ a'_{\mathcal{G}}$  and (ii) for all  $a''' \in A$  with  $a'''_{\mathcal{G}} \succeq a''_{\mathcal{G}}$  it holds that  $M, a''' \models \phi$ . Compare Horty (2001, p. 77). This generalized truth-condition preserves the theorems below that, at first sight, seem to rely on the finiteness condition—Theorems 1(i), 4(i), 4(ii) and 5(i).

$$\begin{array}{ll}
M, a \models p & \text{iff } a \in v(p) \\
M, a \models \neg\phi & \text{iff } M, a \not\models \phi \\
M, a \models \phi \wedge \psi & \text{iff } M, a \models \phi \text{ and } M, a \models \psi \\
M, a \models \diamond\phi & \text{iff there is an } a' \in A \text{ such that } M, a' \models \phi \\
M, a \models [\mathcal{G}]\phi & \text{iff for all } a' \in A \text{ with } a'_G = a_G \text{ it holds that } M, a' \models \phi \\
M, a \models (\mathcal{G})\phi & \text{iff for all } a' \in A \text{ with } a'_G \in \text{Optimal}(\mathcal{G}) \text{ it holds that } M, a' \models \phi.
\end{array}$$

We adopt the following standard conventions: given a deontic game model  $M$ , we write  $M \models \phi$  if for all action profiles  $a$  in  $A$  it holds that  $M, a \models \phi$ . A formula  $\phi$  is *valid* (notation:  $\models \phi$ ) if for all deontic game models  $M$  it holds that  $M \models \phi$ .

### 3.3 Validities and invalidities

We are now in a position to justify the answers to the opening questions of this technical section. To do so, we check the validity of some general claims about relations between group actions and obligations on the one hand and subgroup actions and obligations on the other. But first we make a general observation. Because deontic game models are possible-worlds models in disguise, our agentive and deontic group modalities inherit some basic properties of modalities whose truth-conditions are specified by universal conditions on possible worlds. Accordingly, all of our modalities are *normal* modal operators: they validate the principle of necessitation and the principle of distribution over implication.<sup>13</sup> Although it is straightforward to define non-normal agentive and deontic group modalities on the basis of our normal ones, we avoid unnecessary complications and stick to normal group modalities in the theorems below.<sup>14</sup> All claims are proved in the “Appendix”.

Our first theorem states some of the basic properties of our deontic logic of collective action. First, if group  $\mathcal{G}$  has an obligation to see to it that  $\phi$ , then group  $\mathcal{G}$  has the ability to see to it that  $\phi$ , that is, it is possible that group  $\mathcal{G}$  sees to it that  $\phi$ . This is the well-known principle that ‘ought’ implies ‘can’.<sup>15</sup>

Secondly, group  $\mathcal{G}$  sees to it that both  $\phi$  and  $\psi$  if and only if both group  $\mathcal{G}$  sees to it that  $\phi$  and group  $\mathcal{G}$  sees to it that  $\psi$ . Thirdly, group  $\mathcal{G}$  ought to see to it that both  $\phi$  and  $\psi$  if and only if both group  $\mathcal{G}$  ought to see to it that  $\phi$  and group  $\mathcal{G}$  ought to see to

<sup>13</sup> That is, every modality  $\star$  such that  $\star \in \{\Box, [\mathcal{G}], (\mathcal{G})\}$  validates the principle of necessitation: ‘if  $\models \phi$ , then  $\models \star\phi$ ’, and the principle of distribution over implication: ‘ $\models \star(\phi \rightarrow \psi) \rightarrow (\star\phi \rightarrow \star\psi)$ ’.

<sup>14</sup> We could introduce two additional non-normal operators, ‘Group  $\mathcal{G}$  of agents deliberately sees to it that  $\phi$ ’ (formalized as  $[\mathcal{G}]\phi$ ) and ‘Group  $\mathcal{G}$  of agents deliberately ought to see to it that  $\phi$ ’ (formalized as  $(\mathcal{G})\phi$ ), and stipulate the following truth-conditions for them:  $M, a \models [\mathcal{G}]\phi$  if and only if  $M, a \models [\mathcal{G}]\phi \wedge \neg\Box\phi$ , and  $M, a \models (\mathcal{G})\phi$  if and only if  $M, a \models (\mathcal{G})\phi \wedge \neg\Box\phi$ . The validities for these new operators are the same as the ones listed below, with two exceptions: in Theorems 1(ii) and 1(iii), the condition  $\neg\Box\phi \wedge \neg\Box\psi$  must be added to the left-hand side of the bi-conditional. Horty and Belnap (1995) study the single-agent version of the non-normal agentive modality  $[\mathcal{G}]$ .

<sup>15</sup> Ethical theorists debate whether the principle of ‘ought’ implies ‘can’ holds in every conceivable situation. Key references include Kekes (1984), Sinnott-Armstrong (1988), Howard-Snyder (2006), and Graham (2011).

it that  $\psi$ . Note that the validity of the third claim implies that the statement (2') above,  $(i)(\phi \wedge \psi) \rightarrow ((i)\phi \wedge (i)\psi)$ , is also valid.

**Theorem 1** *Let  $\phi, \psi \in \Omega$  and  $\mathcal{G} \subseteq N$ . Then*

- (i)  $\models (\mathcal{G})\phi \rightarrow \diamond[\mathcal{G}]\phi$
- (ii)  $\models [\mathcal{G}](\phi \wedge \psi) \leftrightarrow ([\mathcal{G}]\phi \wedge [\mathcal{G}]\psi)$
- (iii)  $\models (\mathcal{G})(\phi \wedge \psi) \leftrightarrow ((\mathcal{G})\phi \wedge (\mathcal{G})\psi)$ .

Our second theorem is on relations between group actions and subgroup actions. Let  $\mathcal{F}$  be a subgroup of a group  $\mathcal{G}$ . First, if subgroup  $\mathcal{F}$  sees to it that  $\phi$ , then group  $\mathcal{G}$  sees to it that  $\phi$ .<sup>16</sup> Secondly, the converse implication does not hold: it might be that group  $\mathcal{G}$  sees to it that  $\phi$ , whereas subgroup  $\mathcal{F}$  does not see to it that  $\phi$ . For a counterexample, note that in the deontic game model  $M_1$  of Fig. 1, it holds that  $M_1, (a_i, a_j) \models [i, j]p$  and  $M_1, (a_i, a_j) \not\models [i]p$ . Therefore, the statement (1') above,  $[i, j]\phi \rightarrow ([i]\phi \wedge [j]\phi)$ , is invalid.

**Theorem 2** *Let  $\phi \in \Omega$ . Then*

- (i)  $\models [\mathcal{F}]\phi \rightarrow [\mathcal{G}]\phi$  for all  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$
- (ii)  $\not\models [\mathcal{G}]\phi \rightarrow [\mathcal{F}]\phi$  for some  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$ .

Our third theorem is on relations between group obligations and subgroup obligations. Let  $\mathcal{F}$  be a subgroup of a group  $\mathcal{G}$ . First, the fact that subgroup  $\mathcal{F}$  ought to see to it that  $\phi$  does not imply that group  $\mathcal{G}$  ought to see to it that  $\phi$ : it might be that subgroup  $\mathcal{F}$  has an obligation to see to it that  $\phi$ , whereas group  $\mathcal{G}$  does not have an obligation to see to it that  $\phi$ . Secondly, the converse implication does not hold either: it might be that group  $\mathcal{G}$  has an obligation to see to it that  $\phi$ , whereas subgroup  $\mathcal{F}$  does not have an obligation to see to it that  $\phi$ . For a counterexample to the second claim, note that in the deontic game model  $M_1$  of Fig. 1, for all action profiles  $a$  it holds that  $M_1, a \models (i, j)q$  and  $M_1, a \not\models (i)q$ . Therefore, the statement (3') above,  $(i, j)\phi \rightarrow ((i)\phi \wedge (j)\phi)$ , is invalid.<sup>17</sup>

<sup>16</sup> This seems awkward as long as we do not make the following distinction: '[t]here is more than one thing that one might mean by  $[\mathcal{G}]\phi$ . First, one might mean that the bearers of  $\mathcal{G}$ , without any outside help, guarantee that  $\phi$ , on the basis of a prior simultaneous real choice by each of them. There is also a second, stronger, account. In this version the bearers of  $\mathcal{G}$ , without any outside help, *and with the essential input of each of them*, guarantee that  $\phi$ ' (Belnap et al. 2001, p. 282—notation adapted). We model the first of these two notions of group action.

<sup>17</sup> Regan claims that 'all subgroups of any group which produces best consequences produce best consequences also' and, moreover, that this statement is equivalent to the statement that 'there can be no conflict between the production of good consequences by large groups and by smaller included groups' (Regan 1980, p. 108). Compare Regan (1980, pp. 54 and 56). In the present setting, Regan's first statement is similar to  $(\mathcal{G})\phi \rightarrow (\mathcal{F})\phi$ , which is invalid (Theorem 3(ii)). His second statement is similar to  $(\mathcal{G})\phi \rightarrow \neg(\mathcal{F})\neg\phi$ , which is valid (Theorem 4(ii)). For a detailed comparison between Regan's construction of obligations and one that is closely related to ours, see Horty (2001, Ch. 6).

**Theorem 3** *Let  $\phi \in \Omega$ . Then*

- (i)  $\not\models (\mathcal{F})\phi \rightarrow (\mathcal{G})\phi$  for some  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$
- (ii)  $\not\models (\mathcal{G})\phi \rightarrow (\mathcal{F})\phi$  for some  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$ .

Our fourth theorem is also on relations between group obligations and subgroup obligations. Let  $\mathcal{F}$  be a subgroup of group  $\mathcal{G}$ . First, if subgroup  $\mathcal{F}$  ought to see to it that  $\phi$  is true, then it is not the case that group  $\mathcal{G}$  ought to see to it that  $\phi$  is false. Secondly, if group  $\mathcal{G}$  ought to see to it that  $\phi$  is true, then it is not the case that subgroup  $\mathcal{F}$  ought to see to it that  $\phi$  is false.

**Theorem 4** *Let  $\phi \in \Omega$  and  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$ . Then*

- (i)  $\models (\mathcal{F})\phi \rightarrow \neg(\mathcal{G})\neg\phi$
- (ii)  $\models (\mathcal{G})\phi \rightarrow \neg(\mathcal{F})\neg\phi$ .

Our fifth theorem is also on relations between group obligations and subgroup obligations, but now we consider two subgroups rather than one. To state the theorem clearly, we introduce some additional terminology. A group *has an obligation regarding  $\phi$*  if it ought to see to it that  $\phi$  is true or it ought to see to it that  $\phi$  is false. Two groups *have contradictory obligations regarding  $\phi$* , if the one ought to see to it that  $\phi$  is true, whereas the other ought to see to it that  $\phi$  is false. Two groups *have consistent obligations regarding  $\phi$* , if they do not have contradictory obligations regarding  $\phi$ .

Let  $\mathcal{F}_1$  and  $\mathcal{F}_2$  be subgroups of a group  $\mathcal{G}$ . First, if two subgroups  $\mathcal{F}_1$  and  $\mathcal{F}_2$  have no common members, then their obligations regarding  $\phi$  are consistent. Secondly, if two subgroups  $\mathcal{F}_1$  and  $\mathcal{F}_2$  do have common members, then they might have contradictory obligations regarding  $\phi$ . Thirdly, if two subgroups  $\mathcal{F}_1$  and  $\mathcal{F}_2$  of a group  $\mathcal{G}$  have contradictory obligations regarding  $\phi$ , then group  $\mathcal{G}$  does not have an obligation regarding  $\phi$ . Fourthly, and conversely, if group  $\mathcal{G}$  has an obligation regarding  $\phi$ , then the obligations of its subgroups  $\mathcal{F}_1$  and  $\mathcal{F}_2$  regarding  $\phi$  are consistent.

**Theorem 5** *Let  $\phi \in \Omega$ . Then*

- (i)  $\models (\mathcal{F}_1)\phi \rightarrow \neg(\mathcal{F}_2)\neg\phi$  for all  $\mathcal{F}_1, \mathcal{F}_2 \subseteq N$  such that  $\mathcal{F}_1 \cap \mathcal{F}_2 = \emptyset$
- (ii)  $\not\models (\mathcal{F}_1)\phi \rightarrow \neg(\mathcal{F}_2)\neg\phi$  for some  $\mathcal{F}_1, \mathcal{F}_2 \subseteq N$
- (iii)  $\models ((\mathcal{F}_1)\phi \wedge (\mathcal{F}_2)\neg\phi) \rightarrow \neg(\mathcal{G})\phi \wedge \neg(\mathcal{G})\neg\phi$  for all  $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{G} \subseteq N$
- (iv)  $\models ((\mathcal{G})\phi \vee (\mathcal{G})\neg\phi) \rightarrow \neg((\mathcal{F}_1)\phi \wedge (\mathcal{F}_2)\neg\phi)$  for all  $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{G} \subseteq N$ .

To illustrate Theorem 5(ii), we need some seaside drama: far off the coast, a person is on the verge of drowning. He can be saved by Angela, Bob, and Charlotte, but only if they act together in specific pairs. Angela is an excellent swimmer, Bob a moderate one, and although Charlotte is a weak swimmer, she is the only one who can operate the motorized dinghy anchored nearby. If Angela and Bob work together they will be able to save the drowning person by swimming up to him

immediately: Angela will arrive first, Bob somewhat later, and together they will bring the person back to shore. By contrast, if Bob and Charlotte work together they will be able to save the drowning person by using the dinghy: on arrival, Bob will jump into the water, and together he and Charlotte will manage to pull the person into the dinghy and bring him back to shore. Angela and Bob have a collective obligation to swim. Bob and Charlotte have a collective obligation not to swim. These two collective obligations cannot both be fulfilled. After all, the former requires both Angela and Bob to swim and the latter requires both Bob and Charlotte not to swim. This shows that subgroups that have common members can indeed have contradictory obligations regarding some proposition. Note also that the group consisting of Angela, Bob, and Charlotte has neither a collective obligation to swim nor a collective obligation not to swim.<sup>18</sup>

In sum, our deontic game models can be used to specify truth-conditions for formulas expressing group actions, abilities, obligations, and their interrelations. They thereby make it possible to check the validity of some central claims about relations between group obligations and subgroup obligations. In the next section, we argue that our deontic game models can also be used to make a principled distinction—crucial to our defence of collectivism about collective moral responsibility—between individual obligations and member obligations.

#### 4 Individual obligations and member obligations

If a group has a collective obligation and an individual wishes to assess her obligations, it matters whether the individual is a member of the group or not. To make sense of this simple observation within our formal framework, we make a principled distinction between individual obligations and member obligations and what it means to fulfill such obligations.<sup>19</sup> An *individual obligation* is an obligation that an individual should fulfill regardless of the collective obligations of the groups of which she is a member. We say that an individual agent *fulfills* her individual obligation if and only if she performs one of her optimal individual actions. Likewise, a group fulfills its collective obligation if and only if it performs one of its optimal group actions. (Note that, just like an individual agent, a group of agents, by performing an optimal action, brings about all the states of affairs that it ought to bring about.) If an individual is a member of a group, a *member obligation* towards that group is an obligation that an individual should fulfill in order to help ensure that the group fulfills its collective obligation.<sup>20</sup>

---

<sup>18</sup> See Collins (2013, p. 234) for a similar beach example.

<sup>19</sup> If we read ‘direct wrongdoing’ as not fulfilling an individual obligation and ‘participatory wrongdoing’ as not fulfilling a member obligation, then we do something here that Kutz deems impossible: ‘[w]e will never find a single argument for distinguishing between participatory and direct wrongdoing, for the simple reason that there is no single framework for assessing accountability’ (Kutz 2000a, p. 148).

<sup>20</sup> See Collins (2017) for an informal, but more comprehensive, analysis of ‘membership duties’.

How are we to understand the relation between collective obligations and member obligations? To answer this question, consider a strategic situation represented by a deontic game model. Assume that the individual agents involved are able to communicate and team up. Suppose that some individual agents decide to bundle their strengths and act as a team so as to be better able to promote deontic ideality.<sup>21</sup> The members of the newly formed team will then have to determine which of their options are best, that is, which of the group actions available to them as a team are optimal. To discover these optimal group actions, the members of the team must reason from the standpoint of the team rather than from their individual perspectives. In doing so, they are engaged in *we-reasoning* (Hakli et al. 2010; Tuomela 2013).<sup>22</sup> We take it that the outcome of this reasoning process is a set of optimal group actions. The team fulfills its collective obligation if and only if it performs one of those optimal group actions. Accordingly, the team members will have to design and adopt a *group plan* to coordinate their individual actions such that the composition of their individual actions amounts to an optimal group action (Tamminga and Duijf 2017, Section 3). The group plan thus specifies which individual actions the team members should perform to ensure that the team itself fulfills its collective obligation. Once the team members, aiming to fulfill a collective obligation, have adopted a group plan, a team member fulfills her *member obligation* towards the team if and only if she acts according to plan.

Transposing Bratman's characterization of practical reasoning from an individualistic to a collective perspective, we can think of a group plan as providing 'a filter on options that are potential solutions' (Bratman 1987, p. 35) to a group's coordination problem. Once adopted, a group plan coordinates the actions of the group members: it permits them to act *simultaneously* and *unconditionally*, in the full belief that every group member is acting according to plan. We make the natural assumption that after the group members have adopted a group plan to coordinate their actions, it is common knowledge among them that the plan has been adopted. The formation and adoption of a group plan therefore amounts to forming a collective intention. Group actions regulated by such a plan-based collective intention are known as *shared* or *joint actions*.

If there is a morally significant outcome that requires contributions from several individual agents, then each has an obligation to perform a relevant contributory action. In light of this we assume that when the group members adopt a group plan aimed at realizing a morally significant outcome, the individual members have an obligation to contribute to it.

Given a deontic game model in which a group of agents aims to fulfill its collective obligation, it will be convenient to think of a group plan as a non-empty

<sup>21</sup> Held argues that a random collection of agents might even have a duty to do so: 'the random collection can be held morally responsible for failing to make a decision on which action to take' (Held 1970, p. 479). See also Collins (2013) and Hindriks (2019).

<sup>22</sup> In the theoretical economics literature, the preferred term is *team reasoning* (Bacharach 1999, 2006; Sugden 1993, 2000, 2003). Bacharach (2006, p. 121) writes: '[r]oughly, somebody "team-reasons" if she works out the best feasible combination of actions for all the members of her team, then does her part in it.' Hakli et al. (2010) provide a detailed comparison of we-reasoning and team reasoning.

	$a_j$	$a'_j$
$a_i$	1	0
$a'_i$	0	1

Fig. 2 Not sufficient

	$a_j$	$a'_j$
$a_i$	1	1
$a'_i$	1	0

Fig. 3 Not necessary

subset of the set of the group’s optimal actions. By adopting a group plan, the group members agree to perform one of the group actions in the plan. To give a definition of how such a plan specifies the member obligations of the group members, let us first consider plans that consist of a single optimal group action (this might be just one out of several of the group’s optimal actions). Given a plan that consists of a single optimal group action, it is obvious that if each group member performs her contribution to that single optimal group action, then the group itself performs an optimal group action and thus fulfills its group obligation. We therefore propose the following definition: *given a plan consisting of a single optimal group action, a group member fulfills her member obligation specified by the plan if and only if she performs the action that is her component action of the single optimal group action in that plan.* (Things become more complicated, but also more realistic, if a plan consists of several distinct optimal group actions out of which an arbitrary one is to be performed—more on this at the end of this section.)

Individual obligations should be distinguished clearly from member obligations, because fulfilling an individual obligation is neither *necessary* nor *sufficient* for fulfilling a member obligation. This can be illustrated with the Figs. 2 and 3.<sup>23</sup> To see that fulfilling an individual obligation is *not sufficient* for fulfilling a member obligation, consider the model of Fig. 2. There, it holds that agent  $i$  fulfills her individual obligation if and only if she performs one of the actions  $a_i$  and  $a'_i$ , because both actions are optimal for her. Moreover, given the group plan  $\{(a_i, a_j)\}$ , agent  $i$  fulfills her member obligation specified by that plan if and only if she performs action  $a_i$ . Hence, by performing action  $a'_i$  agent  $i$  fulfills her individual obligation without fulfilling her member obligation. Fulfilling an individual obligation is therefore *not sufficient* for fulfilling a member obligation.

To illustrate this, we return to the seaside, where a person is on the verge of drowning. Suppose now that Angela and Bob will save the drowning person if they both swim. They will also save the drowning person if they both use the motorized dinghy. If one of them swims and the other uses the boat, the person will drown. For Angela, it holds that if she has to act regardless of the collective obligations of the groups of which she is a member, then the two actions that are available to her—swimming and using the boat—are equally good: given that she must act alone, she

<sup>23</sup> For clarity, we suppress the valuation function of the models.

does the best she can by performing either of these actions. Accordingly, were she to perform either of these two individually optimal actions, she would fulfill her individual obligation. Given that Angela and Bob coordinate their actions by adopting the group plan to save the drowning person by swimming, it holds that if Angela were to use the boat, then she would fulfill her individual obligation without fulfilling her member obligation.

To see that fulfilling an individual obligation is *not necessary* for fulfilling a member obligation, consider the model of Fig. 3. There, it holds that agent  $i$  fulfills her individual obligation if and only if she performs action  $a_i$ , because it is her only optimal action. Moreover, given the group plan  $\{(a'_i, a_j)\}$ , agent  $i$  fulfills her member obligation specified by the plan if and only if she performs action  $a'_i$ . Hence, by performing action  $a'_i$  agent  $i$  fulfills her member obligation without fulfilling her individual obligation. Fulfilling an individual obligation is therefore *not necessary* for fulfilling a member obligation. As a consequence, reasons for fulfilling an individual obligation might outweigh reasons for fulfilling a member obligation and vice versa.

To illustrate this, let us slightly change the seaside example. Suppose now that Angela and Bob can each save the drowning person on their own. Angela can perform either of two actions: swimming or not swimming. By swimming she will save the drowning person. Likewise, Bob can perform either of two actions: using the boat or not using the boat. By using the boat he will save the drowning person. For Angela, it holds that if she has to act regardless of the collective obligations of the groups of which she is a member, then the best she can do is to swim. Accordingly, she would fulfill her individual obligation were she to swim. Given that Angela and Bob coordinate their actions by adopting the group plan to save the drowning person by Angela's not swimming and Bob's using the boat, then it holds that if Angela were not to swim, she would fulfill her member obligation without fulfilling her individual obligation.

So far we have only studied member obligations specified by a plan consisting of a *single* optimal group action. This is unnecessarily restrictive and unrealistic. To allow for member obligations specified by plans that consist of several distinct optimal group actions, we propose the following general definition: *given a plan consisting of one or more optimal group actions, a group member fulfills her member obligation specified by the plan if and only if she performs an action that is her component action of one of the optimal group actions in that plan.*

It is important to note that a plan consisting of several distinct optimal group actions does not necessarily guarantee the truth of the conditional 'If each group member fulfills her member obligation specified by the plan, then the group itself fulfills its group obligation'. To see this, let us take a second look at the model of Fig. 2. Consider the plan  $\{(a_i, a_j), (a'_i, a'_j)\}$ . Suppose that agent  $i$  performs action  $a_i$  and that agent  $j$  performs action  $a'_j$ . Then both agent  $i$  and agent  $j$  fulfill their member obligation specified by the plan, although by performing the group action  $(a_i, a'_j)$  the group does not fulfill its group obligation. This plan therefore fails to guarantee the truth of the conditional. It fails to solve the group members' coordination problem. It would therefore be irrational to adopt such a plan. It is a *bad* plan.



A *good* plan—a rational plan—solves the group members' coordination problem, that is, it guarantees the truth of the conditional 'If each group member fulfills her member obligation specified by the plan, then the group itself fulfills its group obligation'.<sup>24</sup> Equivalently, it guarantees the truth of the conditional 'If the group does not fulfill its group obligation, then there is at least one group member who does not fulfill her member obligation specified by the plan'.<sup>25</sup> A bad plan fails to guarantee the truth of these conditionals.

To ensure that a plan solves the group members' coordination problem, the plan has to have a certain structure, namely that whichever group member performs whatever action that is her component action of one of the optimal group actions in the plan, the result is always one of the optimal group actions in the plan. To be precise, a plan should be *closed* under component individual actions.<sup>26</sup>

**Definition 4** (*Closedness*) Let  $M = \langle N, (A_i), d, v \rangle$  be a deontic game model. Let  $\mathcal{G} \subseteq N$  be a group of agents. Let  $B \subseteq A_{\mathcal{G}}$ . Then  $B$  is *closed* if and only if for all  $a_{\mathcal{G}}, a'_{\mathcal{G}} \in B$  and all  $i \in \mathcal{G}$  it holds that  $(a_i, a'_{\mathcal{G}-i}) \in B$ .

A few examples:  $\{(a_i, a'_i)\}$  is closed (as is every set consisting of a single group action). Likewise,  $\{(a_i, a_j), (a_i, a'_j)\}$  is closed. The set  $\{(a_i, a_j), (a'_i, a'_j)\}$  is not, because  $(a_i, a'_i)$  is not in it. The set  $\{(a_i, a_j), (a_i, a'_j), (a'_i, a_j)\}$  is not closed, because  $(a'_i, a'_j)$  is not in it.

A closed plan guarantees the truth of the conditional 'If each group member fulfills her member obligation specified by the plan, then the group itself fulfills its group obligation'. If each group member fulfills her member obligation specified by a closed plan, each group member performs an action that is her component action of one of the optimal group actions in that plan. Because the plan is closed, it must be that this combination of the actions performed by the group members is in the plan. Because the plan only contains optimal group actions, this combination of actions is an optimal group action. Hence, the group members perform an optimal group action. Therefore, if each group member fulfills her member obligation specified by a closed plan, then the group itself fulfills its group obligation. A closed plan thus solves the group members' coordination problem.

Our account of the relation between group obligations, member obligations, and individual obligations has a number of notable features. First, it shows that fulfilling an individual obligation is neither necessary nor sufficient for fulfilling a member obligation. Secondly, if each group member fulfills her member obligation specified

<sup>24</sup> Compare Regan: '[i]f the members of  $\mathcal{G}$  all do their part in the best pattern of behaviour for the members of  $\mathcal{G}$  given the behaviour of non-members, it is clear that the members of  $\mathcal{G}$  produce the best consequences possible as a group' (Regan 1980, p. 138—notation adapted).

<sup>25</sup> This is a possible interpretation of the principle 'If a group act is wrong, at least one of its constituent individual acts is wrong' (Jackson 1987, p. 107). On the basis of an informal counterexample, Jackson argues that this principle is *invalid* (Jackson 1987, pp. 102–103). In our view, Jackson's counterexample does *not* invalidate the principle, because it relies on different concepts of obligations—group obligations, member obligations, individual obligations, conditional individual obligations—that are not systematically distinguished.

<sup>26</sup> Our concept of closedness goes back to Nash (1951, p. 290), where it is called *interchangeability*.

by a *closed* plan, then the group itself fulfills its group obligation. Thirdly, if a group does *not* fulfill its group obligation even though its members adopted a plan to solve their coordination problem, then one of three things has happened: (1) the adopted plan was *not* closed, each group member fulfilled her member obligation specified by the plan, but the combination of the actions performed by the group members was not an optimal group action; (2) the adopted plan was *not* closed, there was at least one group member who did not fulfill her member obligation specified by the plan, and the combination of the actions performed by the group members was not an optimal group action; or (3) the adopted plan *was* closed, there was at least one group member who did not fulfill her member obligation specified by the plan, and the combination of the actions performed by the group members was not an optimal group action. The third possibility supports the counterfactual ‘If each group member had fulfilled her member obligation specified by the plan, then the group itself would have fulfilled its group obligation’, while the second possibility does not. We submit that our account of member obligations helps us to better understand reductionism about collective moral responsibility.

## 5 Reductionism revisited

Our analysis does not support logical reductionism about group obligations and individual obligations. Recall that logical reductionism concerns whether statements about collectives are logically equivalent to conjunctions of statements about individuals. As we showed in Sect. 3.3, group obligations do not imply (and nor are they implied by) individual obligations. Nonetheless, it might seem that our analysis supports logical reductionism about group obligations and *member* obligations: given a closed plan consisting of one or more optimal group actions, it holds that if each group member fulfills her member obligation specified by the plan, then the group itself fulfills its group obligation. Therefore, given that the closedness condition is met, the fulfillment of each member obligation implies the fulfillment of a group obligation. Is this a case of logical reductionism?

First, to meet the demands of logical reductionism, the converse conditional must also be valid. Indeed, Jackson argues that the conditional ‘If a group act is right, *and it is in fact performed*, then each individual constituent act is right’ is valid (Jackson 1987, p. 107). In the present setting, it is not. Given a closed plan consisting of one or more optimal group actions, it might be that the group fulfills its group obligation even though none of the group members fulfills her member obligation specified by the plan. The model of Fig. 2 provides a counterexample. Consider the plan  $\{(a_i, a_j)\}$ . Because it consists of a single optimal group action, it is closed. Suppose that agent  $i$  performs action  $a'_i$  and that agent  $j$  performs action  $a'_j$ . Then the group performs the optimal group action  $(a'_i, a'_j)$  even though neither agent  $i$  nor agent  $j$  fulfills her member obligation specified by the plan.

Secondly, the conditional ‘If each group member fulfills her member obligation specified by the plan, then the group itself fulfills its group obligation’ only holds *under the condition* that the plan is closed. If the plan is *not* closed, the conditional

might be false: given a plan that is *not* closed, it might be that each group member fulfills her member obligation specified by the plan even though the group itself does not fulfill its group obligation. To see this, let us return to Fig. 2, where Angela and Bob will save a drowning person if and only if either (1) they both swim or (2) they both use the motorized dinghy. Given that Angela and Bob adopt the (bad!) group plan to save the drowning person either by swimming or by using the boat (that is, they adopt the group plan  $\{(a_i, a_j), (a'_i, a'_j)\}$ ), it holds that if Angela were to swim ( $a_i$ ) and Bob were to use the boat ( $a'_j$ ), then both Angela and Bob would fulfill their member obligation specified by the plan, but the composition ( $a_i, a'_j$ ) of their individual actions would not amount to an optimal group action, and hence Angela and Bob would not fulfill their collective obligation.

Thirdly, an individualist might point to the fact that member obligations are obligations that individuals have and that this is all we need to vindicate logical reductionism. This response is wide of the mark, because a member obligation is of necessity a member obligation *specified by a plan consisting of one or more optimal group actions*. This means that a member obligation essentially includes a reference to a group. Group obligations are conceptually prior to member obligations, since the group members determine their member obligations by reasoning from the group obligation.<sup>27</sup> As mentioned before, this kind of reasoning is called ‘we-reasoning’ (Tuomela 2013). Kutz acknowledges the role of such reasoning in determining member obligations when he writes that ‘[j]ointly acting agents must reason backwards from the nature of the group act to an understanding of what each should do if the group act is to be achieved’ (Kutz 2000a, pp. 83–84). It is hard to see how this position squares with his individualism.

Fourthly, an individualist might admit that member obligations include a reference to a group, but try to avoid the collectivist consequences of this admission by claiming that this reference is inessential, because member obligations are in fact individual obligations and individual obligations do not include any reference to a group. This response is inadequate for two reasons. First, the proposed identification of member obligations and individual obligations is not enough to meet the demands of individualism, because we showed in Sect. 3.3 that group obligations do not imply (and nor are they implied by) individual obligations. Secondly, we argued in Sect. 4 that fulfilling an individual obligation is neither necessary nor sufficient for fulfilling a member obligation specified by a plan: member obligations are not individual obligations.<sup>28</sup>

<sup>27</sup> Copp, Isaacs, and Lawford-Smith defend closely related positions on the connection between group obligations and member obligations. Copp (2007, p. 374) writes: ‘if a collective has a duty, this is itself a moral consideration that gives at least some members a pro tanto duty to do their part in bringing it about that the collective does its duty’. Isaacs (2011, p. 10) maintains that ‘collective obligation lends shape and form to the obligations of individuals’. Compare Isaacs (2011, pp. 57–58). Lawford-Smith (2012, p. 463) argues that ‘[g]roups have an obligation, which divides into parts in its bearing on the parts of the group, the members’.

<sup>28</sup> Our analysis ties in with Copp’s (2006, 2007) claim that individual obligations might conflict with and even outweigh member obligations. Copp argues that member obligations are *pro tanto* obligations rather than all-things-considered obligations. Accordingly, it might be that a group has an all-things-considered obligation and that the group members have *pro tanto* member obligations. Such a *pro tanto* member obligation might be outweighed by an all-things-considered individual obligation.

We conclude from these four considerations that our analysis does not support logical reductionism about group obligations and member obligations. Because logical reductionism about forward-looking and backward-looking collective moral responsibility is not to be had without logical reductionism about group obligations and member obligations, our analysis also fails to support logical reductionism about collective moral responsibility. All in all, our analysis vindicates collectivism about collective obligations and about collective moral responsibility.

**Acknowledgements** Allard Tamminga gratefully acknowledges financial support from the ERC-2013-CoG Project REINS, No. 616512. The authors thank Jan Broersen, Stephanie Collins, Franz Dietrich, Hein Duijf, John Horty, Barteld Kooi, Holly Lawford-Smith, Jesse Mulder, Frederik Van De Putte, the reviewer of this journal, and the audiences who attended our presentations in Amsterdam, Helsinki, Kyiv, Leiden, London, Regensburg, Saint Petersburg, and Venice for their questions, comments and suggestions.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

### Appendix

**Lemma 1** (Horty 2001, p. 74) *Let  $M = \langle N, (A_i), d, v \rangle$  be a deontic game model. Let  $\mathcal{G} \subseteq N$  be a group of agents. Then*

$$Optimal(\mathcal{G}) \neq \emptyset.$$

*Proof* Assume that  $Optimal(\mathcal{G}) = \emptyset$ . By definition of  $M$ , it holds that  $A_{\mathcal{G}}$  is finite and non-empty. Then for all  $a'_{\mathcal{G}} \in A_{\mathcal{G}}$  there is an  $a_{\mathcal{G}} \in A_{\mathcal{G}}$  such that  $a_{\mathcal{G}} \succ a'_{\mathcal{G}}$ . Because  $A_{\mathcal{G}}$  is finite, there must be a cycle  $a'_{\mathcal{G}} \succ a''_{\mathcal{G}} \succ \dots \succ a''_{\mathcal{G}} \succ a'_{\mathcal{G}} = a''_{\mathcal{G}}$  with all members  $a'_i$  in  $A_{\mathcal{G}}$ . Because  $\succ$  is transitive, it holds that  $a''_{\mathcal{G}} \succ a'_{\mathcal{G}}$ , that is,  $a''_{\mathcal{G}} \succeq a'_{\mathcal{G}}$  and  $a'_{\mathcal{G}} \not\succeq a''_{\mathcal{G}}$ . Contradiction. Therefore,  $Optimal(\mathcal{G}) \neq \emptyset$ .  $\square$

**Lemma 2** (Kooi and Tamminga 2008, p. 9) *Let  $M = \langle N, (A_i), d, v \rangle$  be a deontic game model. Let  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$ . Let  $a_{\mathcal{F}}, a'_{\mathcal{F}} \in A_{\mathcal{F}}$  and let  $a''_{\mathcal{G}-\mathcal{F}} \in A_{\mathcal{G}-\mathcal{F}}$ . Then*

$$If\ a_{\mathcal{F}} \succeq a'_{\mathcal{F}},\ then\ (a_{\mathcal{F}}, a''_{\mathcal{G}-\mathcal{F}}) \succeq (a'_{\mathcal{F}}, a''_{\mathcal{G}-\mathcal{F}}).$$

*Proof* Assume that  $a_{\mathcal{F}} \succeq a'_{\mathcal{F}}$ . Let  $a''_{\mathcal{G}-\mathcal{F}} \in A_{\mathcal{G}-\mathcal{F}}$ . Suppose that  $a'''_{\mathcal{G}} \in A_{-\mathcal{G}}$ . Then  $(a''_{\mathcal{G}-\mathcal{F}}, a'''_{\mathcal{G}}) \in A_{-\mathcal{F}}$ . Hence, by our assumption it must be that  $d(a_{\mathcal{F}}, a''_{\mathcal{G}-\mathcal{F}}, a'''_{\mathcal{G}}) \geq d(a'_{\mathcal{F}}, a''_{\mathcal{G}-\mathcal{F}}, a'''_{\mathcal{G}})$ . Therefore, for all  $a'''_{\mathcal{G}} \in A_{-\mathcal{G}}$  it holds that  $d(a_{\mathcal{F}}, a''_{\mathcal{G}-\mathcal{F}}, a'''_{\mathcal{G}}) \geq d(a'_{\mathcal{F}}, a''_{\mathcal{G}-\mathcal{F}}, a'''_{\mathcal{G}})$ . Hence,  $(a_{\mathcal{F}}, a''_{\mathcal{G}-\mathcal{F}}) \succeq (a'_{\mathcal{F}}, a''_{\mathcal{G}-\mathcal{F}})$ .  $\square$

**Theorem 1** *Let  $\phi, \psi \in \mathfrak{Q}$  and  $\mathcal{G} \subseteq N$ . Then*

- (i)  $\models (\mathcal{G})\phi \rightarrow \Diamond[\mathcal{G}]\phi$
- (ii)  $\models [\mathcal{G}](\phi \wedge \psi) \leftrightarrow ([\mathcal{G}]\phi \wedge [\mathcal{G}]\psi)$

$$(iii) \quad \models (\mathcal{G})(\phi \wedge \psi) \leftrightarrow ((\mathcal{G})\phi \wedge (\mathcal{G})\psi).$$

*Proof* (i) Assume that  $M, a \models (\mathcal{G})\phi$ . Then for all  $a' \in A$  with  $a'_G \in \text{Optimal}(\mathcal{G})$  it holds that  $M, a' \models \phi$ . Because  $\text{Optimal}(\mathcal{G}) \neq \emptyset$ , there is an  $a'_G \in \text{Optimal}(\mathcal{G})$ . Suppose that  $a'' \in A$  is such that  $a''_G = a'_G$ . Then it must be that  $M, a'' \models \phi$ . Hence, for all  $a'' \in A$  with  $a''_G = a'_G$  it holds that  $M, a'' \models \phi$ . Then  $M, a' \models [\mathcal{G}]\phi$ . Hence, there is an  $a \in A$  such that  $M, a' \models [\mathcal{G}]\phi$ . Therefore,  $M, a \models \diamond[\mathcal{G}]\phi$ .

(ii)  $(\Rightarrow)$  Assume that  $M, a \models [\mathcal{G}](\phi \wedge \psi)$ . Then for all  $a' \in A$  with  $a'_G = a_G$  it holds that  $M, a' \models \phi \wedge \psi$ . Then for all  $a' \in A$  with  $a'_G = a_G$  it holds that  $M, a' \models \phi$  and for all  $a' \in A$  with  $a'_G = a_G$  it holds that  $M, a' \models \psi$ . Hence, it must be that  $M, a \models [\mathcal{G}]\phi$  and  $M, a \models [\mathcal{G}]\psi$ . Therefore,  $M, a \models [\mathcal{G}]\phi \wedge [\mathcal{G}]\psi$ .

$(\Leftarrow)$  Assume that  $M, a \models [\mathcal{G}]\phi \wedge [\mathcal{G}]\psi$ . Then  $M, a \models [\mathcal{G}]\phi$  and  $M, a \models [\mathcal{G}]\psi$ . Then for all  $a' \in A$  with  $a'_G = a_G$  it holds that  $M, a' \models \phi$  and for all  $a' \in A$  with  $a'_G = a_G$  it holds that  $M, a' \models \psi$ . Then for all  $a' \in A$  with  $a'_G = a_G$  it holds that  $M, a' \models \phi \wedge \psi$ . Therefore,  $M, a \models [\mathcal{G}](\phi \wedge \psi)$ .

(iii)  $(\Rightarrow)$  Assume that  $M, a \models (\mathcal{G})(\phi \wedge \psi)$ . Then for all  $a' \in A$  with  $a'_G \in \text{Optimal}(\mathcal{G})$  it holds that  $M, a' \models \phi \wedge \psi$ . Then for all  $a' \in A$  with  $a'_G \in \text{Optimal}(\mathcal{G})$  it holds that  $M, a' \models \phi$  and for all  $a' \in A$  with  $a'_G \in \text{Optimal}(\mathcal{G})$  it holds that  $M, a' \models \psi$ . Hence, it must be that  $M, a \models (\mathcal{G})\phi$  and  $M, a \models (\mathcal{G})\psi$ . Therefore,  $M, a \models (\mathcal{G})\phi \wedge (\mathcal{G})\psi$ .

$(\Leftarrow)$  Assume that  $M, a \models (\mathcal{G})\phi \wedge (\mathcal{G})\psi$ . Then  $M, a \models (\mathcal{G})\phi$  and  $M, a \models (\mathcal{G})\psi$ . Then for all  $a' \in A$  with  $a'_G \in \text{Optimal}(\mathcal{G})$  it holds that  $M, a' \models \phi$  and for all  $a' \in A$  with  $a'_G \in \text{Optimal}(\mathcal{G})$  it holds that  $M, a' \models \psi$ . Then for all  $a' \in A$  with  $a'_G \in \text{Optimal}(\mathcal{G})$  it holds that  $M, a' \models \phi \wedge \psi$ . Therefore,  $M, a \models (\mathcal{G})(\phi \wedge \psi)$ .  $\square$

**Theorem 2** *Let  $\phi \in \mathfrak{Q}$ . Then*

- (i)  $\models [\mathcal{F}]\phi \rightarrow [\mathcal{G}]\phi$  for all  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$
- (ii)  $\not\models [\mathcal{G}]\phi \rightarrow [\mathcal{F}]\phi$  for some  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$ .

*Proof* (i) Assume that  $M, a \models [\mathcal{F}]\phi$ . Then for all  $a' \in A$  with  $a'_F = a_F$  it holds that  $M, a' \models \phi$ . Suppose that  $a'' \in A$  is such that  $a''_G = a_G$ . Note that  $a_G$  can be written as  $(a_F, a_{G-F})$ . Then it must be that  $a'' = (a_F, a_{G-F}, a''_{-G})$  and  $a''_F = a_F$ . Then it must be that  $M, a'' \models \phi$ . Hence, for all  $a'' \in A$  with  $a''_G = a_G$  it holds that  $M, a'' \models \phi$ . Therefore,  $M, a \models [\mathcal{G}]\phi$ .

(ii) We define a model  $M = \langle N, (A_i), d, v \rangle$ . Let  $N = \{i, j\}$  be its set of agents. Let  $A_i = \{a_i, a'_i\}$  be the set of actions available to agent  $i$  and let  $A_j = \{a_j, a'_j\}$  be the set of actions available to agent  $j$ . Let the deontic ideality function  $d$  be defined as follows:  $d(a_i, a_j) = d(a_i, a'_j) = d(a'_i, a_j) = d(a'_i, a'_j) = 1$ . Let the valuation function  $v$  be defined as follows:  $v(p) = \{(a_i, a_j)\}$ . See Fig. 4. Then it is easy to see that  $M, (a_i, a_j) \models [i, j]p$  and  $M, (a_i, a_j) \not\models [i]p$ .  $\square$

	$a_j$	$a'_j$
$a_i$	$1/p$	1
$a'_i$	1	1

**Fig. 4**  $[i, j]p$ , but not  $[i]p$

	$a_j$	$a'_j$
$a_i$	$1/p$	$1/p$
$a'_i$	1	0

**Fig. 5**  $(i)p$ , but not  $(i, j)p$

**Theorem 3** *Let  $\phi \in \mathfrak{L}$ . Then*

- (i)  $\not\models (\mathcal{F})\phi \rightarrow (\mathcal{G})\phi$  for some  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$
- (ii)  $\not\models (\mathcal{G})\phi \rightarrow (\mathcal{F})\phi$  for some  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$ .

*Proof* (i) Following Horty (2001, p. 126), we define a model  $M = \langle N, (A_i), d, v \rangle$ . Let  $N = \{i, j\}$  be its set of agents. Let  $A_i = \{a_i, a'_i\}$  be the set of actions available to agent  $i$  and let  $A_j = \{a_j, a'_j\}$  be the set of actions available to agent  $j$ . Let the deontic ideality function  $d$  be defined as follows:  $d(a_i, a_j) = d(a_i, a'_j) = d(a'_i, a_j) = 1$  and  $d(a'_i, a'_j) = 0$ . Then  $Optimal(i) = \{a_i\}$  and  $Optimal(i, j) = \{(a_i, a_j), (a_i, a'_j), (a'_i, a_j)\}$ . Let the valuation function  $v$  be defined as follows:  $v(p) = \{(a_i, a_j), (a_i, a'_j)\}$ . See Fig. 5. Then it is easy to see that for all action profiles  $a$  it holds that  $M, a \models (i)p$  and  $M, a \not\models (i, j)p$ .

(ii) Following Horty (2001, p. 125), we define a model  $M = \langle N, (A_i), d, v \rangle$ . Let  $N = \{i, j\}$  be its set of agents. Let  $A_i = \{a_i, a'_i\}$  be the set of actions available to agent  $i$  and let  $A_j = \{a_j, a'_j\}$  be the set of actions available to agent  $j$ . Let the deontic ideality function  $d$  be defined as follows:  $d(a_i, a_j) = d(a'_i, a'_j) = 1$  and  $d(a_i, a'_j) = d(a'_i, a_j) = 0$ . Then  $Optimal(i) = \{a_i, a'_i\}$  and  $Optimal(i, j) = \{(a_i, a_j), (a'_i, a'_j)\}$ . Let the valuation function  $v$  be defined as follows:  $v(p) = \{(a_i, a_j), (a_i, a'_j), (a'_i, a'_j)\}$ . See Fig. 6. Then it is easy to see that for all action profiles  $a$  it holds that  $M, a \models (i, j)p$  and  $M, a \not\models (i)p$ .  $\square$

**Theorem 4** *Let  $\phi \in \mathfrak{L}$  and  $\mathcal{F} \subseteq \mathcal{G} \subseteq N$ . Then*

- (i)  $\models (\mathcal{F})\phi \rightarrow \neg(\mathcal{G})\neg\phi$
- (ii)  $\models (\mathcal{G})\phi \rightarrow \neg(\mathcal{F})\neg\phi$ .

*Proof* (i) Assume that  $M, a \models (\mathcal{F})\phi$ . Then for all  $a' \in A$  with  $a'_\mathcal{F} \in Optimal(\mathcal{F})$  it holds that  $M, a' \models \phi$ . Suppose that  $M, a \models (\mathcal{G})\neg\phi$ . Then for all  $a' \in A$  with  $a'_\mathcal{G} \in Optimal(\mathcal{G})$  it holds that  $M, a' \models \neg\phi$ . Note that  $Optimal(\mathcal{G}) \neq \emptyset$ . Let  $a' \in A$  be such that  $a'_\mathcal{G} \in Optimal(\mathcal{G})$ . Note that  $a'_\mathcal{G} = (a'_\mathcal{F}, a'_{\mathcal{G}-\mathcal{F}})$  for an  $a'_\mathcal{F} \in A_\mathcal{F}$  and an  $a'_{\mathcal{G}-\mathcal{F}} \in A_{\mathcal{G}-\mathcal{F}}$ . Suppose that  $a'_\mathcal{F} \in Optimal(\mathcal{F})$ . Because  $a'_\mathcal{F} \in Optimal(\mathcal{F})$ , it holds

	$a_j$	$a'_j$
$a_i$	$1/p$	$0/p$
$a'_i$	$0$	$1/p$

Fig. 6  $(i, j)p$ , but not  $(i)p$

	$a_j$	$a'_j$
$a_i$	$1/p$	$1/p$
$a'_i$	$0$	$0$

$a_k$

	$a_j$	$a'_j$
$a_i$	$1$	$0$
$a'_i$	$1$	$0$

$a'_k$

Fig. 7  $(i, k)p$  and  $(j, k)\neg p$

that  $M, a' \models \phi$ . Because  $(a'_F, a'_{G-F}) = a'_G \in \text{Optimal}(\mathcal{G})$ , it holds that  $M, a' \models \neg\phi$ . Contradiction. Hence,  $a'_F \notin \text{Optimal}(\mathcal{F})$ . Then there must be an  $a'' \in A$  with  $a''_F \in \text{Optimal}(\mathcal{F})$  such that  $a''_F \succ a'_F$ . Hence,  $a''_F \succeq a'_F$ . By Lemma 2, it must be that  $(a''_F, a'_{G-F}) \succeq (a'_F, a'_{G-F}) = a'_G$ . Note that for all  $a''' \in A$  with  $a'''_F = a''_F$  it holds that  $M, a''' \models \phi$  and that  $(a''_F, a'_{G-F}) \in A_G$ . Hence,  $(a''_F, a'_{G-F}) \notin \text{Optimal}(\mathcal{G})$ . Then there must be an  $a^* \in A$  with  $a^*_G \in \text{Optimal}(\mathcal{G})$  such that  $a^*_G \succ (a''_F, a'_{G-F})$ . Because  $\succ$  is transitive, it must be that  $a^*_G \succ a'_G$ . Hence,  $a'_G \notin \text{Optimal}(\mathcal{G})$ . Contradiction. Therefore,  $M, a \not\models (\mathcal{G})\neg\phi$ . Hence,  $M, a \models \neg(\mathcal{G})\neg\phi$ .

(ii) Analogous to (i). □

**Theorem 5** Let  $\phi \in \mathfrak{Q}$ . Then

- (i)  $\models (\mathcal{F}_1)\phi \rightarrow \neg(\mathcal{F}_2)\neg\phi$  for all  $\mathcal{F}_1, \mathcal{F}_2 \subseteq N$  such that  $\mathcal{F}_1 \cap \mathcal{F}_2 = \emptyset$
- (ii)  $\not\models (\mathcal{F}_1)\phi \rightarrow \neg(\mathcal{F}_2)\neg\phi$  for some  $\mathcal{F}_1, \mathcal{F}_2 \subseteq N$
- (iii)  $\models ((\mathcal{F}_1)\phi \wedge (\mathcal{F}_2)\neg\phi) \rightarrow (\neg(\mathcal{G})\phi \wedge \neg(\mathcal{G})\neg\phi)$  for all  $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{G} \subseteq N$
- (iv)  $\models ((\mathcal{G})\phi \vee (\mathcal{G})\neg\phi) \rightarrow \neg((\mathcal{F}_1)\phi \wedge (\mathcal{F}_2)\neg\phi)$  for all  $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{G} \subseteq N$ .

*Proof* (i) Following Horty (2001, p. 79), assume that  $M, a \models (\mathcal{F}_1)\phi$ . Then for all  $a' \in A$  with  $a'_{F_1} \in \text{Optimal}(\mathcal{F}_1)$  it holds that  $M, a' \models \phi$ . Because  $\text{Optimal}(\mathcal{F}_1) \neq \emptyset$ , there is an  $a' \in A$  with  $a'_{F_1} \in \text{Optimal}(\mathcal{F}_1)$ . Suppose that  $M, a \models (\mathcal{F}_2)\neg\phi$ . Then for all  $a'' \in A$  with  $a''_{F_2} \in \text{Optimal}(\mathcal{F}_2)$  it holds that  $M, a'' \models \neg\phi$ . Because  $\text{Optimal}(\mathcal{F}_2) \neq \emptyset$ , there is an  $a'' \in A$  with  $a''_{F_2} \in \text{Optimal}(\mathcal{F}_2)$ . Because  $\mathcal{F}_1 \cap \mathcal{F}_2 = \emptyset$ , it holds that  $(a'_{F_1}, a''_{F_2}) \in A_{\mathcal{F}_1 \cup \mathcal{F}_2}$ . Let  $a^* = (a'_{F_1}, a''_{F_2}, a'''_{-(\mathcal{F}_1 \cup \mathcal{F}_2)})$  for some  $a''' \in A$  with  $a'''_{-(\mathcal{F}_1 \cup \mathcal{F}_2)} \in A_{-(\mathcal{F}_1 \cup \mathcal{F}_2)}$ . Then  $M, a^* \models \phi$  and  $M, a^* \models \neg\phi$ . Contradiction. Therefore,  $M, a \models \neg(\mathcal{F}_2)\neg\phi$ .

(ii) Following Kooi and Tamminga (2008, p. 16), we define a model  $M = \langle N, (A_i), d, v \rangle$ . Let  $N = \{i, j, k\}$  be its set of agents. Let  $A_i = \{a_i, a'_i\}$  be the set of actions available to agent  $i$ , let  $A_j = \{a_j, a'_j\}$  be the set of actions available to agent  $j$ , and let  $A_k = \{a_k, a'_k\}$  be the set of actions available to agent  $k$ . Let the deontic ideality function  $d$  be defined as follows:  $d(a_i, a_j, a_k) = d(a_i, a_j, a'_k) = d(a_i, a'_j, a_k) = d(a'_i, a_j, a'_k) = 1$  and  $d(a_i, a'_j, a'_k) = d(a'_i, a_j, a_k) = d(a'_i, a'_j, a_k) =$

$d(a'_i, a'_j, a'_k) = 0$ . Let the valuation function  $v$  be defined as follows:  $v(p) = \{(a_i, a_j, a_k), (a_i, a'_j, a_k)\}$ . See Fig. 7. Then it is easy to see that for all action profiles  $a$  it holds that  $M, a \models (i, k)p$  and  $M, a \models (j, k)\neg p$ .

(iii) Use Theorem 4 and contraposition.

(iv) Use (iii) and contraposition. □

## References

- Bacharach, M. (1999). Interactive team reasoning: A contribution to the theory of cooperation. *Research in Economics*, 53, 117–147.
- Bacharach, M. (2006). *Beyond individual choice*. Princeton, NJ: Princeton University Press.
- Belnap, N., Perloff, M., & Xu, M. (2001). *Facing the future*. New York: Oxford University Press.
- Bratman, M. (1987). *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.
- Bratman, M. (1992). Shared cooperative activity. *Philosophical Review*, 101, 327–341.
- Carmo, J. (2010). Collective agency, direct action and dynamic operators. *Logic Journal of the IGPL*, 18, 66–98.
- Chant, S. R. (2007). Unintentional collective action. *Philosophical Explorations*, 10, 245–256.
- Collins, S. (2013). Collectives' duties and collectivization duties. *Australasian Journal of Philosophy*, 91, 231–248.
- Collins, S. (2017). Duties of group agents and group members. *Journal of Social Philosophy*, 48, 38–57.
- Copp, D. (2006). On the agency of certain collective entities: An argument from 'normative autonomy'. *Midwest Studies in Philosophy*, 30, 194–221.
- Copp, D. (2007). The collective moral autonomy thesis. *Journal of Social Philosophy*, 38, 369–388.
- Feinberg, J. (1968). Collective responsibility. *Journal of Philosophy*, 65, 674–688.
- Gilbert, M. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy*, 15, 1–14.
- Gilbert, M. (2006). Who's to blame? Collective moral responsibility and its implications for group members. *Midwest Studies in Philosophy*, 30, 94–114.
- Graham, P. (2011). 'Ought' and ability. *Philosophical Review*, 120, 337–382.
- Hakli, R., Miller, K., & Tuomela, R. (2010). Two kinds of we-reasoning. *Economics and Philosophy*, 26, 291–320.
- Hansson, S. O. (1986). Individuals and collective actions. *Theoria*, 52, 87–97.
- Held, V. (1970). Can a random collection of individuals be morally responsible? *Journal of Philosophy*, 67, 471–481.
- Hilpinen, R. (Ed.). (1971). *Deontic logic*. Dordrecht: D. Reidel Publishing Company.
- Hindriks, F. (2013). The location problem in social ontology. *Synthese*, 190, 413–437.
- Hindriks, F. (2019). The duty to join forces: When individuals lack control. *The Monist*, 102 (in press).
- Horty, J. (1996). Agency and obligation. *Synthese*, 108, 269–307.
- Horty, J. (2001). *Agency and deontic logic*. New York: Oxford University Press.
- Horty, J., & Belnap, N. (1995). The deliberative stit: A study of action, omission, ability, and obligation. *Journal of Philosophical Logic*, 24, 583–644.
- Howard-Snyder, F. (2006). 'Cannot' implies 'not ought'. *Philosophical Studies*, 130, 233–246.
- Isaacs, T. (2011). *Moral responsibility in collective contexts*. New York: Oxford University Press.
- Jackson, F. (1987). Group morality. In P. Pettit, R. Sylvan, & J. Norman (Eds.), *Metaphysics and morality* (pp. 91–110). Oxford: Basil Blackwell.
- Kanger, S. (1957). New foundations for ethical theory. Privately distributed manuscript, first published in R. Hilpinen (Ed.), *Deontic logic* (pp. 36–58). Dordrecht 1971: D. Reidel Publishing Company.
- Kanger, S. (1972). Law and logic. *Theoria*, 38, 105–132.
- Kekes, J. (1984). 'Ought implies can' and two kinds of morality. *Philosophical Quarterly*, 34, 459–467.
- Kooi, B., & Tamminga, A. (2008). Moral conflicts between groups of agents. *Journal of Philosophical Logic*, 37, 1–21.
- Kutz, C. (2000a). *Complicity*. Cambridge: Cambridge University Press.
- Kutz, C. (2000b). Acting together. *Philosophy and Phenomenological Research*, 61, 1–31.



- Lawford-Smith, H. (2012). The feasibility of collectives' actions. *Australasian Journal of Philosophy*, 90, 453–467.
- List, C., & Pettit, P. (2011). *Group agency*. Oxford: Oxford University Press.
- Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, 54, 286–295.
- Osborne, M., & Rubinstein, A. (1994). *A course in game theory*. Cambridge, MA: The MIT Press.
- Pörn, I. (1970). *The logic of power*. Oxford: Basil Blackwell.
- Regan, D. (1980). *Utilitarianism and co-operation*. New York: Oxford University Press.
- Schelling, T. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
- Searle, J. (1990). Collective intentions and actions. In P. R. Cohen, J. Morgan, & M. E. Pollock (Eds.), *Intentions in communication* (pp. 401–415). Cambridge, MA: The MIT Press.
- Sinnott-Armstrong, W. (1988). *Moral dilemmas*. Oxford: Basil Blackwell.
- Sugden, R. (1993). Thinking as a team: Towards an explanation of non-selfish behavior. *Social Philosophy and Policy*, 10, 69–89.
- Sugden, R. (2000). Team preferences. *Economics and Philosophy*, 16, 175–204.
- Sugden, R. (2003). The logic of team reasoning. *Philosophical Explorations*, 6, 165–181.
- Tamminga, A. (2013). Deontic logic for strategic games. *Erkenntnis*, 78, 183–200.
- Tamminga, A., & Duijf, H. (2017). Collective obligations, group plans and individual actions. *Economics and Philosophy*, 33, 187–214.
- Tuomela, R. (1989). Collective action, supervenience, and constitution. *Synthese*, 80, 243–266.
- Tuomela, R. (2013). *Social ontology*. New York: Oxford University Press.
- Tuomela, R., & Miller, K. (1988). We-intentions. *Philosophical Studies*, 53, 367–389.
- von Kutschera, F. (1986). Bewirken. *Erkenntnis*, 24, 253–281.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.