



An impossibility result on methodological individualism

Hein Duijf¹ · Allard Tamminga² · Frederik Van De Putte^{3,4}

Accepted: 10 April 2021 / Published online: 17 May 2021
© The Author(s) 2021

Abstract Methodological individualists often claim that any social phenomenon can ultimately be explained in terms of the actions and interactions of individuals. Any Nagelian version of methodological individualism requires that there be bridge laws that translate social statements into individualistic ones. We show that Nagelian individualism can be put to logical scrutiny by making the relevant social and individualistic languages fully explicit and mathematically precise. In particular, we prove that the social statement that a group of (at least two) agents performs a deontically admissible group action cannot be expressed in a well-established deontic logic of agency that models every combination of actions, omissions, abilities, and obligations of finitely many individual agents.

Keywords Methodological individualism · Impossibility result · Collective admissibility · Modal logic · Expressivity · Bisimulation

✉ Hein Duijf
h.w.a.duijf@vu.nl

¹ Department of Philosophy, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

² Department of Philosophy, University of Greifswald, Baderstraße 6, 17489 Greifswald, Germany

³ Department of Philosophy, Ghent University, Blandijnberg 2, 9000 Ghent, Belgium

⁴ Erasmus Institute for Philosophy and Economics, Erasmus School of Philosophy, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands

1 Introduction

Methodological individualists often claim that any social phenomenon can ultimately be explained in terms of the actions and interactions of individuals.¹ This reductionist thesis does not simply follow from the associated ontological thesis, according to which every social entity ultimately consists of individual entities and their interrelations. To make a strong case for the problematic character of methodological individualism, we examine the requirements for individualistic explanations of social phenomena under the assumption that ontological individualism is true. To be specific, in this paper we assume that groups are sets of individual agents and that group actions are sets of individual actions.

An influential version of methodological individualism relies on Ernest Nagel's (1961, Ch. 11) classical model of intertheoretical reduction to specify the requirements for individualistic explanations of social phenomena.² We use the term 'Nagelian individualism' to refer to any variety of methodological individualism that uses Nagel's model to give an account of individualistic explanations of social phenomena. According to Nagelian individualism, any individualist explanation of a (possibly probabilistic) law about social phenomena must consist in deriving this social law from a set of (possibly probabilistic) laws about actions and interactions of individuals plus some *bridge laws* that translate social statements containing social terms into individualistic statements that contain only individualistic terms. Therefore, bridge laws that translate social statements into individualistic ones are essential for Nagelian individualism.³

What characteristics should such a bridge law have? First, a bridge law shows that a particular social statement can be expressed in a specific individualistic language. Whether or not there is such a bridge law therefore depends on the *expressive power* of the specific individualistic language that is supposed to do the translating. Secondly, a bridge law cannot be a stipulative definition, because the social statement that requires translation already has an extension. To specify a

¹ Elster (1989, p. 13) concurs: "The elementary unit of social life is the individual human action. To explain social institutions and social change is to show how they arise as the result of the action and interaction of individuals. This view, often referred to as methodological individualism, is in my view trivially true." According to Elster (1985, p. 5), methodological individualism is "the doctrine that all social phenomena—their structure and their change—are in principle explicable in ways that only involve individuals [...]. Methodological individualism thus conceived is a form of reductionism."

² Although methodological individualism has some precursors in neoclassical economics, the doctrine was first introduced as a methodological maxim for the social sciences by Max Weber. A first major philosophical debate on individualism started in the 1950s, with key contributions from Agassi, Jarvie, Popper, and Watkins; a second one began in the 1980s, with key contributions from Boudon, Coleman, Elster, and Roemer. These debates generated several distinct versions of methodological individualism. (See Udehn (2001) for a historical review and detailed references.) More recently, the philosophy of explanation entered the debate on individualism (see for instance Kincaid (1986), Bhargava (1992), Sawyer (2002, 2003), List and Spiekermann (2013)). Our paper falls within this last tradition, where Nagelian individualism is of central importance.

³ The claim that social statements can always be translated into individualistic statements is what List and Pettit (2011, p. 3) call 'eliminativism': "Anything ascribed to a group [...] can be re-expressed by reference to its members."

bridge law for the social statement, we have to find an individualistic statement such that the social statement and the individualistic statement are logically or nomologically co-extensive.⁴ (Two statements are logically co-extensive if and only if their extensions are the same in all possible worlds. Two statements are nomologically co-extensive if and only if their extensions are the same in all possible worlds where the relevant background laws hold.) A necessary condition for a bridge law is that the social statement to be translated must be *logically* or at least *nomologically equivalent* to its individualistic translation: they must be true under exactly the same circumstances, where the range of relevant circumstances might be restricted by *individualistic* background laws.⁵

These philosophical considerations on the expressive power of individualistic languages in the social sciences can be made precise by adopting a logical point of view. In formal logic, several techniques have been developed to study the relative expressive power of formal languages. We therefore use formal logic to investigate whether or not there is a bridge law that translates particular social statements into a specific individualistic language \mathcal{Q}_i . Adapting the definition of expressive power that is standard in modal logic to our current needs, we obtain the following: an individualistic language \mathcal{Q}_i is *at least as expressive as* a social language \mathcal{Q} if and only if for every statement χ in \mathcal{Q} there is a statement ϕ in \mathcal{Q}_i such that ϕ and χ are logically equivalent.⁶ Accordingly, if we wish to prove that there is no bridge law that translates a particular social statement χ into a specific individualistic language \mathcal{Q}_i , it suffices to show the following: there are no individualistic statements ϕ and ψ in \mathcal{Q}_i such that the individualistic statement ϕ and the social statement χ are nomologically equivalent modulo the individualistic statement ψ , where ψ can be thought of as a finite conjunction of individualistic background laws. (By substituting a tautology for ψ , it follows from this that there is no individualistic statement ϕ in \mathcal{Q}_i such that the individualistic statement ϕ and the social statement χ are logically equivalent.)

The social statements that we focus on in this paper are *collective deontic admissibility statements* of the form “Group \mathcal{G} of agents performs a deontically admissible group action”. The notion of collective admissibility is central to formal studies of forward-looking and backward-looking collective moral responsibility and of collective rationality.⁷ It is also an indispensable notion in team-reasoning

⁴ This requirement is also endorsed by Kincaid (1986, p. 494): “[R]eduction requires lawlike co-extensionality between the primitive predicates of social theory and some predicate in the reducing theory. [...] Reduction does not require equivalence of meaning, but [...] it does require biconditional bridge laws connecting primitive terminology of social theories with terminology of individualist theory.”

⁵ If methodological individualists are not to compromise their position, they must presuppose that all of the restricting background laws can be expressed in the individualistic language.

⁶ This is the definition that is needed to assess Quinton’s (1975, p. 23) claim that “every statement about a social object is equivalent to and can be replaced by a statement in which only individuals are referred to and in which the predicates, whether the same as or different from those of the original statement, are predicates of individuals”.

⁷ Horty (2001, p. 130) relies on collective deontic admissibility to define group obligations, and Tamminga and Duijf use collective deontic admissibility to analyse collective rationality

and we-reasoning accounts of cooperation.⁸ The main technical objective of this paper is to establish an *impossibility result* on collective deontic admissibility statements: there are no bridge laws translating collective deontic admissibility statements into an individualistic formal language that is widely used to model actions, omissions, abilities, and obligations of finitely many individual agents. More specifically, the social statement that a group performs a deontically admissible group action cannot be translated into the individualistic language that is built from a necessity operator, operators for individual agency, and operators for individual deontic admissibility.

Taking a reflective step back, does our impossibility result *refute* Nagelian individualism? Obviously, it does not: there might be individualistic languages other than the one used in this paper that allow for a bridge law translating collective deontic admissibility statements. Our contribution to the debate is therefore *methodological*. Where the philosophical literature on methodological individualism mainly consists in clarifying its meaning and scope, we show that by making the social and the individualistic languages fully explicit and mathematically precise, we can precisely determine those situations in which Nagelian individualism is tenable and those in which it is not.

Our paper is organized as follows. In Sect. 2 we define the social language Ω (of which the individualistic language Ω_i is a sublanguage) and interpret the sentences of that language by way of *deontic game models*. The thus defined deontic logic of agency is then used to specify the problem of finding bridge laws that translate collective deontic admissibility statements into the individualistic language Ω_i (Sect. 3). In the next two sections we prove that there are no such bridge laws. In Sect. 4—a technical section that may be skipped by those who are more interested in results than proofs—we introduce the notion of individualistic bisimulation and define two constructions that transform a given deontic game model into deontic game models that validate exactly the same individualistic formulas as the given model but give different truth values to collective deontic admissibility statements. In Sect. 5, we establish our impossibility result: there are no bridge laws that translate collective deontic admissibility statements into the individualistic language Ω_i . Lastly, in Sect. 6, we formulate two research questions that are prompted by our impossibility result.

Footnote 7 continued

(2017, pp. 200–201) and backward-looking collective moral responsibility (2017, § 6). Constants for collective deontic admissibility were first introduced by Tamminga et al. (forthcoming).

⁸ Bacharach's (1999, 2006) and Sugden's (1993, 2000, 2003) accounts of 'team-reasoning' are closely related to Tuomela's (2013) study of 'we-reasoning'. Bacharach (2006, p. 121) writes: "Roughly, 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."

2 Language and semantics

The logic of agency for which we prove our impossibility theorems is a logic in the tradition of *stit* ('sees to it that') logics of agency.⁹ These logics—and especially their semantics—have been central to the formal study of actions, omissions, abilities, and obligations. We first introduce the social language \mathfrak{Q} and its individualistic sublanguage \mathfrak{Q}_i (Sect. 2.1) and then use what we call *deontic game models* to specify truth-conditions for the formulas in these languages (Sect. 2.2).

2.1 A social language and its individualistic sublanguage

The social language \mathfrak{Q} and its individualistic sublanguage \mathfrak{Q}_i are built from a countable set \mathfrak{P} of atomic formulas and a finite set \mathcal{N} of individual agents. Throughout the paper, we use p and q as variables for atomic formulas, i and j as variables for individual agents, and \mathcal{F} and \mathcal{G} as variables for sets of individual agents. The *social language* \mathfrak{Q} is the smallest set S (in terms of set-theoretical inclusion) that satisfies conditions (i) through (vii) below. The *individualistic language* \mathfrak{Q}_i is the smallest set S that satisfies conditions (i) through (vi).

- | | |
|---|----------------------------|
| (i) $\mathfrak{P} \subseteq S$ | (Atomic formulas) |
| (ii) If $\phi \in S$, then $\neg\phi \in S$ | (Negation) |
| (iii) If $\phi \in S$ and $\psi \in S$, then $(\phi \wedge \psi) \in S$ | (Conjunction) |
| (iv) If $\phi \in S$, then $\Box\phi \in S$ | (Necessity) |
| (v) If $i \in \mathcal{N}$ and $\phi \in S$, then $[i]\phi \in S$ | (Individual agency) |
| (vi) If $i \in \mathcal{N}$, then $\star_{\{i\}} \in S$ | (Individual admissibility) |
| (vii) If $\mathcal{G} \subseteq \mathcal{N}$, then $\star_{\mathcal{G}} \in S$. | (Collective admissibility) |

Note that condition (vi) is a special case of condition (vii). The individualistic language \mathfrak{Q}_i is therefore the sublanguage of \mathfrak{Q} that consists of exactly all formulas in \mathfrak{Q} that do not contain collective admissibility formulas. The operators \vee , \rightarrow , \leftrightarrow , and \diamond abbreviate the usual constructions. Parentheses and braces are left out if the omission does not give rise to ambiguities. Accordingly, the formulas $\phi \wedge \psi$ and \star_i are shorthand for the formulas $(\phi \wedge \psi)$ and $\star_{\{i\}}$, respectively.

The social language \mathfrak{Q} can be used to express *necessity statements* like “It is settled true that ϕ ” (formalized as $\Box\phi$), *individual agentive statements* like “Agent i sees to it that ϕ ” (formalized as $[i]\phi$), *individual deontic admissibility statements* like “Agent i performs a deontically admissible individual action” (formalized as \star_i), and *collective deontic admissibility statements* like “Group \mathcal{G} of agents performs a deontically admissible group action” (formalized as $\star_{\mathcal{G}}$).

⁹ Seminal works include Kanger (1957), Pörn (1970), von Kutschera (1986), Horty and Belnap (1995), and Horty (1996). See Belnap et al. (2001) and Horty (2001) for textbook presentations of these logics. Their connections to game theory have been explored by Kooi and Tamminga (2008), Tamminga (2013), Van De Putte et al. (2017), Duijf (2018, Chs. 1 and 4) and Tamminga and Hindriks (2020).

The individualistic language \mathcal{Q}_i has been central to formal studies of omissions, abilities, and obligations of individual agents.¹⁰ *Individual omission statements* like “Agent i refrains from seeing to it that ϕ ” can be formalized as $[i]\neg[i]\phi$, *individual ability statements* like “Agent i is able to see to it that ϕ ” can be formalized as $\diamond[i]\phi$, and *individual obligation statements* like “Agent i ought to see to it that ϕ ” can be formalized as $\Box(\star_i \rightarrow \phi)$.

2.2 Deontic game models

Truth-conditions for the formulas in the social language \mathcal{Q} (and hence for the formulas in the individualistic language \mathcal{Q}_i) are specified in terms of *deontic game models* (see Tamminga and Hindriks (2020)). These models represent the individual actions and the group actions that are available to a finite set of individual agents and evaluate the deontic admissibility of these actions. Deontic game models combine possible-worlds semantics for standard deontic logic with game forms from elementary game theory. A model for standard deontic logic consists of a non-empty set of possible worlds, a non-empty set of deontically ideal possible worlds (which is a subset of the set of possible worlds), and a valuation function that assigns to each atomic formula a set of possible worlds where that formula is true. A game form from elementary game theory specifies which individual actions are available to which individual agents at a given moment in time. Every possible combination of individual actions (one for each individual agent) is a possible action profile. Basically, a deontic game model is a possible-worlds model for standard deontic logic where the possible worlds are replaced by possible action profiles. Hence, a deontic game model consists of a non-empty set of possible action profiles, a non-empty set of deontically ideal possible action profiles (which 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 \mathcal{N} of individual agents. Each agent i in \mathcal{N} is assigned a non-empty and finite set A_i of available individual actions. We use a_i , b_i , and c_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 $\times_{i \in \mathcal{N}} A_i$ of all the individual agents’ sets of actions. Note that A is non-empty, because all the A_i ’s are non-empty. We use a , b , and c as variables for action profiles in A .¹¹ The set of deontically ideal action profiles is defined by a deontic ideality function d that assigns to each 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).¹² We require that there

¹⁰ Horty and Belnap (1995) present formal analyses of omissions and abilities, including more refined versions of the ones we present here. Horty’s (2001) analysis of what an individual agent ought to do is the same as the one we present here. Tamminga et al. (forthcoming) note that Horty’s formalization of individual obligations can be given an Andersonian-Kangerian definition in terms of necessity and individual deontic admissibility.

¹¹ We adopt the notational conventions of Osborne and Rubinstein (1994, Sect. 1.7).

¹² 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

be at least one deontically ideal action profile in A . Finally, a valuation function v assigns to each atomic formula p in \mathfrak{P} a (possibly empty) set $v(p)$ of action profiles:¹³

Definition 1 (Deontic Game Model) A *deontic game model* M is a quadruple $\langle \mathcal{N}, (A_i), d, v \rangle$, where \mathcal{N} is a finite set of individual agents, for each agent i in \mathcal{N} it holds that A_i is a non-empty and finite set of actions available to agent i , $d : A \rightarrow \{0, 1\}$ is a deontic ideality function such that there is at least one a in A with $d(a) = 1$, and $v : \mathfrak{P} \rightarrow \wp(A)$ is a valuation function.

As an illustration, consider the two-agent deontic game model M_1 with individual agents i and j . Let $A_i = \{a_i, b_i, c_i\}$ and $A_j = \{a_j, b_j, c_j\}$. The set A of action profiles is $A_i \times A_j$. Let $d(a_i, b_j) = d(a_i, c_j) = d(b_i, a_j) = d(b_i, b_j) = d(c_i, a_j) = 1$ and let $d(a_i, a_j) = d(b_i, c_j) = d(c_i, b_j) = d(c_i, c_j) = 0$. Lastly, let $v(p) = \{(a_i, a_j)\}$ and $v(q) = \{(b_i, a_j), (b_i, b_j), (b_i, c_j)\}$ (the valuations of the other atomic formulas are left unspecified). M_1 can be pictured as in Fig. 1.

Let us now turn to group actions and their admissibility. As stated in the introduction, we make the simplifying ontological assumption that groups are sets of individual agents and that group actions are sets of individual actions.¹⁴ Accordingly, for each set $\mathcal{G} \subseteq \mathcal{N}$ of individual agents, the set $A_{\mathcal{G}}$ of group actions that are available to \mathcal{G} is defined as the Cartesian product $\times_{i \in \mathcal{G}} A_i$. We use $a_{\mathcal{G}}$ and $b_{\mathcal{G}}$ as variables for group actions in $A_{\mathcal{G}}$. Moreover, if $a_{\mathcal{G}}$ is a group action of group \mathcal{G} and if $\mathcal{F} \subseteq \mathcal{G}$, we use $a_{\mathcal{F}}$ to denote the subgroup action that is \mathcal{F} 's component subgroup action of the group action $a_{\mathcal{G}}$. Finally, we use $-\mathcal{G}$ to denote the relative complement $\mathcal{N} - \mathcal{G}$.

To specify truth-conditions for individual and collective deontic admissibility statements, we order the group actions that are available to a group \mathcal{G} by way of a dominance relation. Intuitively, group action $a_{\mathcal{G}}$ weakly dominates group action $b_{\mathcal{G}}$ in deontic game model M (notation: $a_{\mathcal{G}} \succeq_M b_{\mathcal{G}}$) if and only if $a_{\mathcal{G}}$ promotes deontic ideality in M at least as well as $b_{\mathcal{G}}$, regardless of what the group's non-members do:

Definition 2 (Dominance) Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model. Let $\mathcal{G} \subseteq \mathcal{N}$ be a set of individual agents. Let $a_{\mathcal{G}}, b_{\mathcal{G}} \in A_{\mathcal{G}}$. Then

$$a_{\mathcal{G}} \succeq_M b_{\mathcal{G}} \text{ iff for all } c_{-\mathcal{G}} \in A_{-\mathcal{G}} \text{ it holds that } d(a_{\mathcal{G}}, c_{-\mathcal{G}}) \geq d(b_{\mathcal{G}}, c_{-\mathcal{G}}).$$

Footnote 12 continued

action profiles unanimously as 'morally acceptable' or 'morally unacceptable' (McNamara, 1996, p. 163).

¹³ Deontic game models closely resemble Schelling's (1960, p. 84) *pure-collaboration games* and Bacharach's (2006, p. 122) *coordination contexts*. As a counterpart to zero-sum games, Schelling (1960, p. 84) examines cooperation with pure-collaboration games "in which the players win or lose together, having identical preferences regarding the outcome."

¹⁴ Collective intentionality is therefore missing in our account of group actions. In a formal framework that is comparable to the one used in the current paper, Tamminga and Duijf (2017, § 3) introduce 'group plans' to model Bratman's (2014) plan-based collective intentions.

	a_j	b_j	c_j
a_i	0/p	1	1
b_i	1/q	1/q	0/q
c_i	1	0	0

Fig. 1 Deontic game model M_1

Strong dominance is defined in terms of weak dominance: $a_G \succ_M b_G$ if and only if $a_G \succeq_M b_G$ and $b_G \not\preceq_M a_G$.

We use this dominance relation to define the sets of deontically admissible actions that are available to individual agents and groups of agents. (Individual agents are here thought of as singleton groups.) A group action is *deontically admissible* in deontic game model M if there is no other available group action in M that strongly dominates it.¹⁵

Definition 3 (Deontic Admissibility) Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model. Let $\mathcal{G} \subseteq \mathcal{N}$ be a set of individual agents. Then the set of \mathcal{G} 's *deontically admissible actions* in M , denoted by $\text{Adm}_M(\mathcal{G})$, is given by $\text{Adm}_M(\mathcal{G}) = \{a_G \in A_G : \text{there is no } b_G \in A_G \text{ such that } b_G \succ_M a_G\}$.

To illustrate deontic admissibility, consider the deontic game model M_1 of Fig. 1 and note that for M_1 it holds that $\text{Adm}_{M_1}(i) = \{a_i, b_i\}$, $\text{Adm}_{M_1}(j) = \{a_j, b_j\}$, and $\text{Adm}_{M_1}(i, j) = \{(a_i, b_j), (a_i, c_j), (b_i, a_j), (b_i, b_j), (c_i, a_j)\}$.

For every deontic game model M , it holds that $\text{Adm}_M(\mathcal{G})$ is non-empty if \mathcal{G} is non-empty.¹⁶ Given a deontic game model and an action profile, a group of agents performs a deontically admissible group action if and only if the group's contribution to the action profile is a deontically admissible group action. To be precise, given a deontic game model M and an action profile a , a group \mathcal{G} performs a deontically admissible action if and only if a_G is one of \mathcal{G} 's deontically admissible group actions in M .

Lastly, how does this formal framework model individual agency? According to the *stit* theory of agency, by performing an action an individual agent restricts the total set of possible situations at a given moment in time to those possible situations that are consistent with her action. An individual agent sees to it that a particular statement is true if this statement is true in every possible situation that is consistent with her action. To transpose these ideas to the current framework, we identify the total set of possible situations at a given moment in time with the set A of possible

¹⁵ On admissibility, see Arrow (1951, p. 429), Luce & Raiffa (1957, pp. 287 and 307), Savage (1972, p. 21) and Kohlberg & Mertens (1986, § 2.7.A). Admissibility sets are central to Horty's (2001, Ch. 4) formalization of what he calls *dominance act utilitarianism*.

¹⁶ Note that for the grand coalition \mathcal{N} we always have $\text{Adm}_M(\mathcal{N}) = \{a \in A : d(a) = 1\}$.

action profiles in a deontic game model. Given a deontic game model and an action profile, an individual agent sees to it that a particular statement is true if and only if the agent’s contribution to the action profile suffices to ensure the truth of this statement, regardless of the actions of the other individual agents. To be precise, given a deontic game model M and an action profile a in A , an individual agent i sees to it that ϕ if and only if for all action profiles b with $b_i = a_i$ it holds that ϕ is true at b .

In line with the above considerations, the truth-conditions for the formulas in the language \mathfrak{Q} are the following:

Definition 4 (Truth-Conditions) Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model. Let $i \in \mathcal{N}$ be an individual agent and let $\mathcal{G} \subseteq \mathcal{N}$ be a set of individual agents. Let $a \in A$ be an action profile. Let $p \in \mathfrak{P}$ be an atomic formula and $\phi, \psi \in \mathfrak{Q}$ be arbitrary formulas. Then

- $(M, a) \models p$ iff $a \in v(p)$
- $(M, a) \models \star_{\mathcal{G}}$ iff $a_{\mathcal{G}} \in \text{Adm}_M(\mathcal{G})$
- $(M, a) \models \neg\phi$ iff $(M, a) \not\models \phi$
- $(M, a) \models \phi \wedge \psi$ iff $(M, a) \models \phi$ and $(M, a) \models \psi$
- $(M, a) \models \Box\phi$ iff for all $b \in A$ it holds that $(M, b) \models \phi$
- $(M, a) \models [i]\phi$ iff for all $b \in A$ with $b_i = a_i$ it holds that $(M, b) \models \phi$.

To illustrate these truth-conditions, consider the deontic game model M_1 of Fig. 1 and note that at action profile (a_i, a_j) of M_1 it holds that p , although neither agent i nor agent j sees to it that p : we have $(M_1, (a_i, a_j)) \models p \wedge \neg[i]p \wedge \neg[j]p$. Moreover, agent i is able to see to it that q , although agent j is not: we have $(M_1, (a_i, a_j)) \models \Diamond[i]q \wedge \neg\Diamond[j]q$. Lastly, because $a_i \in \text{Adm}_{M_1}(i)$, $a_j \in \text{Adm}_{M_1}(j)$, and $(a_i, a_j) \notin \text{Adm}_{M_1}(i, j)$, at action profile (a_i, a_j) of M_1 it holds that agent i and agent j both perform a deontically optimal individual action, although the group $\{i, j\}$ of agents does not do so: we have $(M_1, (a_i, a_j)) \models \star_i \wedge \star_j \wedge \neg\star_{\{i, j\}}$.

In general, 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 ϕ is *valid* (notation: $\models \phi$) if for all deontic game models M it holds that $M \models \phi$. Any ordered pair (M, a) that consists of a deontic game model M and one of its action profiles a is a *pointed* deontic game model.

3 When are there no bridge laws?

In the introduction, we saw that Nagelian individualism requires that there be bridge laws that translate social statements into individualistic ones. Recall that to specify a bridge law for a particular social statement, we have to find an individualistic statement such that the social statement and the individualistic statement are

logically or nomologically equivalent: they must be true under exactly the same circumstances, where the range of relevant circumstances might be restricted by individualistic background laws.

The social statements we focus on in this paper are collective deontic admissibility statements of the form “Group \mathcal{G} of agents performs a deontically admissible group action” (formalized as $\star_{\mathcal{G}}$). Are there bridge laws that translate collective deontic admissibility statements into the individualistic language \mathfrak{Q}_i ? That is, is there an individualistic statement ϕ in \mathfrak{Q}_i and a finite conjunction ψ of individualistic background laws in \mathfrak{Q}_i such that ϕ and $\star_{\mathcal{G}}$ are equivalent in all circumstances where the background laws ψ hold?¹⁷ More precisely, are there individualistic statements ϕ and ψ in \mathfrak{Q}_i such that ϕ and $\star_{\mathcal{G}}$ are equivalent modulo ψ , that is, such that $\models \psi \rightarrow (\phi \leftrightarrow \star_{\mathcal{G}})$?

Let us illustrate the depth of the problem by returning to the individualistic statement $\star_i \wedge \star_j$ and the social statement $\star_{\{i,j\}}$. It is easy to see that they are not equivalent: in deontic game model M_1 we have $(M_1, (a_i, a_j)) \models \star_i \wedge \star_j \wedge \neg \star_{\{i,j\}}$. This does not suffice, however, to prove that there is no bridge law that translates $\star_{\{i,j\}}$ into \mathfrak{Q}_i . Although this simple observation does show that $\star_i \wedge \star_j$ and $\star_{\{i,j\}}$ are not *logically* equivalent, it does not show that they are not *nomologically* equivalent. After all, there might still be a (non-contradictory) individualistic statement ψ such that $\star_i \wedge \star_j$ and $\star_{\{i,j\}}$ are equivalent modulo ψ , that is, such that $\models \psi \rightarrow ((\star_i \wedge \star_j) \leftrightarrow \star_{\{i,j\}})$.¹⁸ To rule out this possibility, we must show that for every non-contradictory individualistic statement ψ in \mathfrak{Q}_i it holds that $\not\models \psi \rightarrow ((\star_i \wedge \star_j) \leftrightarrow \star_{\{i,j\}})$. And even if we were to show that $\star_i \wedge \star_j$ and $\star_{\{i,j\}}$ are not nomologically equivalent, we would not have proved that there is no bridge law that translates $\star_{\{i,j\}}$ into \mathfrak{Q}_i . To prove that there is no such bridge law, we would also have to show that every individualistic statement that differs from $\star_i \wedge \star_j$ is not nomologically equivalent to $\star_{\{i,j\}}$ either. How can we prove that there is no such bridge law without surveying all conceivable candidates?

In the next two sections, we adapt the notion of bisimulation from modal logic and use it to prove that for every group \mathcal{G} of two or more members, it holds that (i) there are no individualistic statements ϕ and ψ in \mathfrak{Q}_i such that ϕ and $\star_{\mathcal{G}}$ are equivalent modulo ψ , unless ψ is a contradiction (Theorem 4), and (ii) there is no individualistic statement ϕ in \mathfrak{Q}_i such that ϕ and $\star_{\mathcal{G}}$ are equivalent (Corollary 1). Consequently, there are no bridge laws that translate collective deontic admissibility statements into the individualistic language \mathfrak{Q}_i . This is our impossibility result on Nagelian individualism.

¹⁷ Our impossibility result can easily be generalized to infinite sets of individualistic background laws. We stick to the finite case for expository reasons only.

¹⁸ We exclude the trivial case where the finite conjunction ψ of background laws is a contradiction, because from a contradiction anything follows.

4 Individualistic bisimilarity and two constructions

To prove our impossibility result, we rely on a notion of similarity between deontic game models. The central consideration is that if two deontic game models are similar in the relevant sense, then they validate exactly the same individualistic formulas, or equivalently, then there is no individualistic formula that is true in the one model and false in the other. All lemmas and theorems in this section are proved in the “Appendix”.

4.1 Individualistic bisimilarity

In modal logic several notions of model similarity have been developed to capture the structural conditions models must satisfy to validate exactly the same formulas. One prominent such notion is bisimulation.¹⁹ We adapt the standard notion of bisimulation and introduce the notion of *individualistic bisimulation* (Definition 5). After having done so, we prove that pointed deontic game models that are individualistically bisimilar validate exactly the same individualistic formulas (Theorem 1).

Definition 5 (Individualistic Bisimulation) Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ and $M' = \langle \mathcal{N}, (A'_i), d', v' \rangle$ be deontic game models. Let $A = \times_{i \in \mathcal{N}} A_i$ and $A' = \times_{i \in \mathcal{N}} A'_i$. A relation $R \subseteq A \times A'$ is an ι -bisimulation between M and M' if for all $a \in A$ and $a' \in A'$ with $(a, a') \in R$ it holds that

- (i) for all $p \in \mathfrak{P}$ it holds that $a \in v(p)$ iff $a' \in v'(p)$
- (ii) for all $i \in \mathcal{N}$ it holds that $a_i \in \text{Adm}_M(i)$ iff $a'_i \in \text{Adm}_{M'}(i)$
- (iii) for all $b \in A$ there is a $b' \in A'$ such that $(b, b') \in R$
- (iv) for all $b' \in A'$ there is a $b \in A$ such that $(b, b') \in R$
- (v) for all $i \in \mathcal{N}$ and $b \in A$: if $a_i = b_i$, then there is a $b' \in A'$ such that $a'_i = b'_i$ and $(b, b') \in R$
- (vi) for all $i \in \mathcal{N}$ and $b' \in A'$: if $a'_i = b'_i$, then there is a $b \in A$ such that $a_i = b_i$ and $(b, b') \in R$.

We write $(M, a) \rightleftharpoons_{\iota} (M', a')$ if there is an ι -bisimulation relation R between M and M' such that $(a, a') \in R$. Moreover, we write $(M, a) \equiv_{\mathfrak{Q}_i} (M', a')$ if (M, a) and (M', a') validate exactly the same individualistic formulas, that is, if for all $\phi \in \mathfrak{Q}_i$ it holds that $(M, a) \models \phi$ iff $(M', a') \models \phi$.

We can now state our first theorem: if two pointed deontic game models are ι -bisimilar, they validate exactly the same individualistic formulas.

Theorem 1 (ι -Bisimulation Theorem) *For all pointed deontic game models (M, a) and (M', a') : if $(M, a) \rightleftharpoons_{\iota} (M', a')$, then $(M, a) \equiv_{\mathfrak{Q}_i} (M', a')$.*

¹⁹ Blackburn et al. (2001, § 2.2) provide a textbook presentation of bisimulation in modal logic.

4.2 Two constructions

Given a deontic action model M and a set $\mathcal{G} \subseteq \mathcal{N}$ of individual agents, we construct two deontic action models, $M'_{\mathcal{G}}$ and $M''_{\mathcal{G}}$, and prove that they are ι -bisimilar to M . Consequently, they validate exactly the same individualistic formulas. The models $M'_{\mathcal{G}}$ and $M''_{\mathcal{G}}$ take center stage in the proof of our impossibility result.

To specify the models $M'_{\mathcal{G}}$ and $M''_{\mathcal{G}}$, we need some additional notation. We use x and y as variables for the elements of $\{+, -\}^{\mathcal{N}}$. Following the notational conventions for action profiles, we use x_i to denote the projection of x onto i , and we use $x_{\mathcal{G}}$ to denote the projection of x onto \mathcal{G} . Accordingly, we can write the $2n$ -tuple $(a_1, x_1, \dots, a_n, x_n)$ as an ordered pair (a, x) of two n -tuples, where $a = (a_1, \dots, a_n)$ and $x = (x_1, \dots, x_n)$. Given a set $\mathcal{G} \subseteq \mathcal{N}$ of agents and an $x \in \{+, -\}^{\mathcal{N}}$, we say that $x_{\mathcal{G}}$ is *constant* if for every $i, j \in \mathcal{G}$ it holds that $x_i = x_j$.

4.2.1 The unit transform

Given a set \mathcal{G} of agents, we unit-transform any given deontic game model M into a new deontic game model $M'_{\mathcal{G}}$ by first duplicating the individual actions that are available to the individual agents in the given model. This results in new sets A'_i of available individual actions, one for each individual agent i . The new set A' of action profiles is defined as the Cartesian product $\times_{i \in \mathcal{N}} A'_i$. The new deontic ideality function d' copies the 0s of the given model but tinkers with its 1s (depending on whether $x_{\mathcal{G}}$ is constant). The new valuation function v' copies the valuation function of the given model.

Given a deontic game model M and a set \mathcal{G} of agents, we first give a precise definition of the unit \mathcal{G} -transformation of the deontic game model M into the deontic game model $M'_{\mathcal{G}}$. Secondly, we prove that M and $M'_{\mathcal{G}}$ are ι -bisimilar.

Definition 6 Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model and let $\mathcal{G} \subseteq \mathcal{N}$. Then the unit \mathcal{G} -transform of M , denoted by $M'_{\mathcal{G}} = \langle \mathcal{N}, (A'_i), d', v' \rangle$, is given by

$$A'_i = A_i \times \{+, -\} \text{ for every individual agent } i \in \mathcal{N}$$

$$d'(a, x) = \begin{cases} d(a), & \text{if } x_{\mathcal{G}} \text{ is constant} \\ 0, & \text{otherwise.} \end{cases}$$

$$(a, x) \in v'(p) \text{ iff } a \in v(p).$$

It is easy to check that $M'_{\mathcal{G}}$ is a deontic game model.

As an illustration of this transformation, we consider the unit $\{i, j\}$ -transform of deontic game model M_1 from Fig. 1. The resulting deontic game model $(M_1)'_{\{i, j\}}$ is pictured in Fig. 2. To prove that a deontic game model and its unit \mathcal{G} -transform are

	$(a_j, +)$	$(a_j, -)$	$(b_j, +)$	$(b_j, -)$	$(c_j, +)$	$(c_j, -)$
$(a_i, +)$	0/p	0/p	1	0	1	0
$(a_i, -)$	0/p	0/p	0	1	0	1
$(b_i, +)$	1/q	0/q	1/q	0/q	0/q	0/q
$(b_i, -)$	0/q	1/q	0/q	1/q	0/q	0/q
$(c_i, +)$	1	0	0	0	0	0
$(c_i, -)$	0	1	0	0	0	0

Fig. 2 Deontic game model $(M_1)_{\{i,j\}}$

ι -bisimilar, it is essential that the ι -bisimulation relation connect the admissible individual actions of both models appropriately. It may therefore be helpful to note that $\text{Adm}_{(M_1)_{\{i,j\}}}(i) = \{(a_i, +), (a_i, -), (b_i, +), (b_i, -)\}$ and $\text{Adm}_{(M_1)_{\{i,j\}}}(j) = \{(a_j, +), (a_j, -), (b_j, +), (b_j, -)\}$.

The transformation of M into M'_G preserves dominance relations between individual actions in the following way:

Lemma 1 *Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model, let $\mathcal{G} \subseteq \mathcal{N}$, and let $M'_G = \langle \mathcal{N}, (A'_i), d', v' \rangle$ be the unit \mathcal{G} -transform of M . Let $a_i, b_i \in A_i$ and $x_i, y_i \in \{+, -\}$. Then*

- (i) $a_i \succeq_M b_i$ iff $(a_i, x_i) \succeq_{M'_G} (b_i, x_i)$
- (ii) If $(a_i, x_i) \succ_{M'_G} (b_i, y_i)$, then $a_i \succ_M b_i$.

With Lemma 1 in hand, we can prove that M and its unit \mathcal{G} -transform M'_G are individually bisimilar:

Theorem 2 *Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model, let $\mathcal{G} \subseteq \mathcal{N}$, and let $M'_G = \langle \mathcal{N}, (A'_i), d', v' \rangle$ be the unit \mathcal{G} -transform of M . Let $R_1 \subseteq A \times A'$ be given by $\{(a, (a, x)) : a \in A \text{ and } x \in \{+, -\}^{\mathcal{N}}\}$. Then R_1 is an ι -bisimulation between M and M'_G .*

4.2.2 The zero transform

The second operation on deontic game models we define is identical to the previous one, except for its definition of the deontic ideality function. Given a set \mathcal{G} of agents, we zero-transform any given deontic game model M into a new deontic game model $M''_{\mathcal{G}}$ by first duplicating the individual actions that are available to the individual agents in the given model. This results in new sets A''_i of available individual actions, one for each individual agent i . The new set A'' of action profiles is defined as the Cartesian product $\times_{i \in \mathcal{N}} A''_i$. The new deontic ideality function d'' copies the 1s

	$(a_j, +)$	$(a_j, -)$	$(b_j, +)$	$(b_j, -)$	$(c_j, +)$	$(c_j, -)$
$(a_i, +)$	0/p	1/p	1	1	1	1
$(a_i, -)$	1/p	0/p	1	1	1	1
$(b_i, +)$	1/q	1/q	1/q	1/q	0/q	1/q
$(b_i, -)$	1/q	1/q	1/q	1/q	1/q	0/q
$(c_i, +)$	1	1	0	1	0	1
$(c_i, -)$	1	1	1	0	1	0

Fig. 3 Deontic game model $(M_1)''_{\{i,j\}}$

of the given model but tinkers with its 0s (depending on whether x_G is constant). The new valuation function v'' copies the valuation function of the given model.

Given a deontic game model M and a set G of agents, we first give a precise definition of the zero G -transformation of the deontic game model M into the deontic game model M''_G . Secondly, we prove that M and M''_G are ι -bisimilar.

Definition 7 Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model and let $G \subseteq \mathcal{N}$. Then the zero G -transform of M , denoted by $M''_G = \langle \mathcal{N}, (A''_i), d'', v'' \rangle$, is given by

$$A''_i = A_i \times \{+, -\} \text{ for every individual agent } i \in \mathcal{N}$$

$$d''(a, x) = \begin{cases} d(a), & \text{if } x_G \text{ is constant} \\ 1, & \text{otherwise.} \end{cases}$$

$$(a, x) \in v''(p) \text{ iff } a \in v(p).$$

It is easy to check that M''_G is a deontic game model.

As an illustration of this transformation, we consider the zero $\{i, j\}$ -transform of deontic game model M_1 from Fig. 1. The resulting deontic game model $(M_1)''_{\{i,j\}}$ is pictured in Fig. 3. To prove that a deontic game model and its zero G -transform are ι -bisimilar, it is essential that the ι -bisimulation relation connect the admissible individual actions of both models appropriately. It may therefore be helpful to note that $\text{Adm}_{(M_1)''_{\{i,j\}}}(i) = \{(a_i, +), (a_i, -), (b_i, +), (b_i, -)\}$ and $\text{Adm}_{(M_1)''_{\{i,j\}}}(j) = \{(a_j, +), (a_j, -), (b_j, +), (b_j, -)\}$.

The transformation of M into M''_G preserves dominance relations between individual actions in the same way as in Lemma 1:

Lemma 2 Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model, let $G \subseteq \mathcal{N}$, and let $M''_G = \langle \mathcal{N}, (A''_i), d'', v'' \rangle$ be the zero G -transform of M . Let $a_i, b_i \in A_i$ and $x_i, y_i \in \{+, -\}$. Then

- (i) $a_i \succeq_M b_i$ iff $(a_i, x_i) \succeq_{M''_G} (b_i, x_i)$

- (ii) If $(a_i, x_i) \succ_{M''_{\mathcal{G}}} (b_i, y_i)$, then $a_i \succ_M b_i$.

Finally, using Lemma 2, we show that M and its zero \mathcal{G} -transform $M''_{\mathcal{G}}$ are individually bisimilar:

Theorem 3 Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model, let $\mathcal{G} \subseteq \mathcal{N}$, and let $M''_{\mathcal{G}} = \langle \mathcal{N}, (A''_i), d'', v'' \rangle$ be the zero \mathcal{G} -transform of M . Let $R_2 \subseteq A \times A''$ be given by $\{(a, (a, x)) : a \in A \text{ and } x \in \{+, -\}^{\mathcal{N}}\}$. Then R_2 is an ι -bisimulation between M and $M''_{\mathcal{G}}$.

5 An impossibility result

We now establish our impossibility result: there are no individualistic statements ϕ and ψ in \mathfrak{L}_i such that ϕ and $\star_{\mathcal{G}}$ are equivalent modulo ψ , unless ψ is a contradiction (Theorem 4). As a consequence, there is no ϕ in \mathfrak{L}_i such that ϕ and $\star_{\mathcal{G}}$ are equivalent (Corollary 1).

Theorem 4 Let $\mathcal{G} \subseteq \mathcal{N}$ be a non-empty and non-singleton group of agents. Then there are no $\phi, \psi \in \mathfrak{L}_i$ such that $\not\models \neg\psi$ and $\models \psi \rightarrow (\phi \leftrightarrow \star_{\mathcal{G}})$.

Proof We argue by contradiction. Suppose there are $\phi, \psi \in \mathfrak{L}_i$ such that $\not\models \neg\psi$ and $\models \psi \rightarrow (\phi \leftrightarrow \star_{\mathcal{G}})$. Because $\not\models \neg\psi$, there is a deontic game model $M = \langle \mathcal{N}, (A_i), d, v \rangle$ and an action profile $a \in A$ such that $(M, a) \models \psi$. Because $\models \psi \rightarrow (\phi \leftrightarrow \star_{\mathcal{G}})$, it holds that $(M, a) \models \phi \leftrightarrow \star_{\mathcal{G}}$. Then either (i) $(M, a) \models \phi$ and $(M, a) \models \star_{\mathcal{G}}$, or (ii) $(M, a) \not\models \phi$ and $(M, a) \not\models \star_{\mathcal{G}}$. We show that each of these cases leads to a contradiction.

- (i) Suppose $(M, a) \models \phi$ and $(M, a) \models \star_{\mathcal{G}}$. Let $M'_{\mathcal{G}} = \langle \mathcal{N}, (A'_i), d', v' \rangle$ be the unit \mathcal{G} -transform of M . Let $x \in \{+, -\}^{\mathcal{N}}$ and $i, j \in \mathcal{G}$ be such that $x_i \neq x_j$. By Theorem 2, it holds that $(M, a) \rightleftharpoons_i (M'_{\mathcal{G}}, (a, x))$. By Theorem 1, we have $(M'_{\mathcal{G}}, (a, x)) \models \psi$ and $(M'_{\mathcal{G}}, (a, x)) \models \phi$. Because $\models \psi \rightarrow (\phi \leftrightarrow \star_{\mathcal{G}})$, it holds that $(M'_{\mathcal{G}}, (a, x)) \models \star_{\mathcal{G}}$. By definition of d' and because $x_{\mathcal{G}}$ is not constant, for all $c'_{-\mathcal{G}} \in A'$ it holds that $d'(a_{\mathcal{G}}, x_{\mathcal{G}}, c'_{-\mathcal{G}}) = 0$. Because $M'_{\mathcal{G}}$ is a deontic game model, there is a $b' \in A'$ such that $d'(b') = 1$. Hence, $(a_{\mathcal{G}}, x_{\mathcal{G}}) \notin \text{Adm}_{M'_{\mathcal{G}}}(\mathcal{G})$. Hence, $(M'_{\mathcal{G}}, (a, x)) \not\models \star_{\mathcal{G}}$. Contradiction.
- (ii) Suppose $(M, a) \not\models \phi$ and $(M, a) \not\models \star_{\mathcal{G}}$. Let $M''_{\mathcal{G}} = \langle \mathcal{N}, (A''_i), d'', v'' \rangle$ be the zero \mathcal{G} -transform of M . Let $x \in \{+, -\}^{\mathcal{N}}$ and $i, j \in \mathcal{G}$ such that $x_i \neq x_j$. By Theorem 3, it holds that $(M, a) \rightleftharpoons_i (M''_{\mathcal{G}}, (a, x))$. By Theorem 1, we have $(M''_{\mathcal{G}}, (a, x)) \models \psi$ and $(M''_{\mathcal{G}}, (a, x)) \not\models \phi$. Because $\models \psi \rightarrow (\phi \leftrightarrow \star_{\mathcal{G}})$, it holds that $(M''_{\mathcal{G}}, (a, x)) \not\models \star_{\mathcal{G}}$. By definition of d'' and because $x_{\mathcal{G}}$ is not constant, for all $c''_{-\mathcal{G}} \in A''_{\mathcal{G}}$ it holds that $d''(a_{\mathcal{G}}, x_{\mathcal{G}}, c''_{-\mathcal{G}}) = 1$. Hence, $(a_{\mathcal{G}}, x_{\mathcal{G}}) \in \text{Adm}_{M''_{\mathcal{G}}}(\mathcal{G})$. Hence, $(M''_{\mathcal{G}}, (a, x)) \models \star_{\mathcal{G}}$. Contradiction.

Because each of these cases leads to a contradiction, we conclude that there are no $\phi, \psi \in \mathfrak{Q}_i$ such that $\not\models \neg\psi$ and $\models \psi \rightarrow (\phi \leftrightarrow \star_G)$. \square

By substituting a tautology for ψ , it follows from Theorem 4 that there is no ϕ in \mathfrak{Q}_i such that ϕ and \star_G are equivalent:

Corollary 1 *Let $\mathcal{G} \subseteq \mathcal{N}$ be a non-empty and non-singleton group of agents. Then there is no $\phi \in \mathfrak{Q}_i$ such that $\models \phi \leftrightarrow \star_G$.*

Therefore, there are no bridge laws that translate collective deontic admissibility statements into the individualistic language \mathfrak{Q}_i .

6 Future research

Impossibility results trigger new research questions. We address two of them. The first one is: are there individualistic languages more expressive than \mathfrak{Q}_i that do allow for a translation of collective deontic admissibility? To vindicate methodological individualism (that is, to answer this question positively), it would be necessary to define a new individualistic language and a model theory to give truth-conditions for the formulas of this language, and to show that there are bridge laws that translate collective deontic admissibility statements into this new individualistic language. Of course, collective deontic admissibility statements are not the only social statements that can be studied formally.

To formulate the second research question, note first that Theorem 4 entails that there are deontic game models in which the formula $(\bigwedge_{i \in \mathcal{G}} \star_i) \rightarrow \star_G$ is false.²⁰ Accordingly, for any group of individual agents there are situations in which each group member performs a deontically admissible individual action even though the combination of these individual actions does not amount to a deontically admissible group action.²¹ Nonetheless, there are deontic game models in which the formula $(\bigwedge_{i \in \mathcal{G}} \star_i) \rightarrow \star_G$ is true: consider, for instance, deontic game models with exactly one deontically ideal action profile (there are also other instances). In such models the relation between individual deontic admissibility and collective deontic admissibility is straightforward: if every group member performs a deontically admissible individual action, then the group itself performs a deontically admissible group action. The second research question can now be formulated as follows: do all the deontic game models that validate this implication have a specific structural property in common? More specifically, is there a model-theoretic property that

²⁰ Suppose $\models (\bigwedge_{i \in \mathcal{G}} \star_i) \rightarrow \star_G$. Let $\phi = p \vee \neg p$ and $\psi = \bigwedge_{i \in \mathcal{G}} \star_i$. It is easy to see that $\phi, \psi \in \mathfrak{Q}_i$ and $\not\models \neg\psi$ and $\models \psi \rightarrow (\phi \leftrightarrow \star_G)$. This contradicts Theorem 4. Therefore, $\not\models (\bigwedge_{i \in \mathcal{G}} \star_i) \rightarrow \star_G$.

²¹ Likewise, Theorem 4 entails that there are deontic game models in which the formula $(\bigwedge_{i \in \mathcal{G}} \neg \star_i) \rightarrow \neg \star_G$ is false. Accordingly, for any group of individual agents there are situations in which no group member performs a deontically admissible individual action even though the combination of these individual actions does amount to a deontically admissible group action.

defines the class \mathcal{C} of deontic game models for which it holds that $M \in \mathcal{C}$ if and only if $M \models (\bigwedge_{i \in \mathcal{G}} \star_i) \rightarrow \star_{\mathcal{G}}$?²² After such a structural property has been specified, we can also study it from a dynamic perspective: it might be asked which operations transform an arbitrary deontic game model into a model that has this property.²³ These and related questions offer a new take on the systematic study of cooperation in particular, and of methodological individualism in general.

Acknowledgements Hein Duijf gratefully acknowledges funding from the ERC-2017-CoG Project SEA, No. 771074. Hein Duijf and Allard Tamminga gratefully acknowledge funding from the ERC-2013-CoG project REINS, No. 616512. Frederik Van De Putte's research was funded by a Marie Skłodowska-Curie Fellowship, No. 795329, by a grant from the Research Foundation—Flanders (FWO-Vlaanderen), No. 12Q1918N, and by a grant from the Dutch Research Council, No. VI.Vidi.191.105. The authors thank Frank Hindriks, Hannes Leitgeb, Jeanne Peijnenburg, Friedrich Reinmuth, Olivier Roy, the reviewer of this journal, and the audiences who attended our presentations at the University of Bayreuth, Ghent University, the University of Greifswald, the University of Groningen, the University of Manchester, the Università Cattolica del Sacro Cuore (Milan), Saint Louis University, the University of Twente, and Utrecht University for their questions, comments and suggestions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix: proofs

Theorem 1 (*1-Bisimulation Theorem*) *For all pointed deontic game models (M, a) and (M', a') : if $(M, a) \rightleftharpoons_i (M', a')$, then $(M, a) \equiv_{\Omega_i} (M', a')$.*

Proof By structural induction on ϕ .

Basis: Suppose $(M, a) \rightleftharpoons_i (M', a')$. We check cases p and \star_j .

p : By Definition 5(i), for all $p \in \mathfrak{P}$ it holds that $a \in v(p)$ iff $a' \in v'(p)$. Hence, for all $p \in \mathfrak{P}$ it holds that $(M, a) \models p$ iff $(M', a') \models p$.

²² This is closely related to Van Benthem's (1984) notion of a modal formula that characterizes a frame property. See also Blackburn et al. (2001, p. 126). Note that Theorem 4 entails that the class \mathcal{C} cannot be characterized by an individualistic formula in Ω_i .

²³ Tamminga and Duijf (2017) study how the public adoption of a group plan changes the context in which agents make a decision about what to do. They show that after a deontic game model is updated with what they call an "optimal and interchangeable group plan", it holds that if every group member performs a deontically admissible individual action in the changed decision context, then the group itself performs a deontically admissible group action.

★_i: By Definition 5(ii), for all $i \in \mathcal{N}$ it holds that $a_i \in \text{Adm}_M(i)$ iff $a'_i \in \text{Adm}_{M'}(i)$. Hence, for all $i \in \mathcal{N}$ it holds that $(M, a) \models \star_i$ iff $(M', a') \models \star_i$.

Induction hypothesis For all pointed deontic game models (M_1, a_1) and (M'_1, a'_1) and all $\psi \in \mathfrak{L}_i$ with fewer operators than ϕ : if $(M_1, a_1) \rightleftharpoons_i (M'_1, a'_1)$, then $(M_1, a_1) \models \psi$ iff $(M'_1, a'_1) \models \psi$.

Induction step Suppose $(M, a) \rightleftharpoons_i (M', a')$. We check case $[i]$. The other cases are proved analogously.

$[i]$: Suppose $(M, a) \models [i]\psi$. Suppose $b' \in A'$ such that $a'_i = b'_i$. By Definition 5(vi), there is a $b \in A$ such that $a_i = b_i$ and $(b, b') \in R$. Then $(M, b) \rightleftharpoons_i (M', b')$. Because $(M, a) \models [i]\psi$ and $a_i = b_i$, we have $(M, b) \models \psi$. By the Induction Hypothesis, $(M', b') \models \psi$. Hence, for all $b' \in A'$ such that $a'_i = b'_i$ it holds that $(M', b') \models \psi$. Hence, $(M', a') \models [i]\psi$. Suppose $(M', a') \models [i]\psi$. Suppose $b \in A$ such that $a_i = b_i$. By Definition 5(v), there is a $b' \in A'$ such that $a'_i = b'_i$ and $(b, b') \in R$. Then $(M, b) \rightleftharpoons_i (M', b')$. Because $(M', a') \models [i]\psi$ and $a'_i = b'_i$, we have $(M', b') \models \psi$. By the Induction Hypothesis, $(M, b) \models \psi$. Hence, for all $b \in A$ such that $a_i = b_i$ it holds that $(M, b) \models \psi$. Hence, $(M, a) \models [i]\psi$.

Therefore, $(M, a) \equiv_{\mathfrak{L}_i} (M', a')$. □

Lemma 1 Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model, let $\mathcal{G} \subseteq \mathcal{N}$, and let $M'_{\mathcal{G}} = \langle \mathcal{N}, (A'_i), d', v' \rangle$ be the unit \mathcal{G} -transform of M . Let $a_i, b_i \in A_i$ and $x_i, y_i \in \{+, -\}$. Then

- (i) $a_i \succeq_M b_i$ iff $(a_i, x_i) \succeq_{M'_{\mathcal{G}}} (b_i, x_i)$
- (ii) If $(a_i, x_i) \succ_{M'_{\mathcal{G}}} (b_i, y_i)$, then $a_i \succ_M b_i$.

Proof (i) (\Rightarrow) Suppose $a_i \succeq_M b_i$. Consider an arbitrary $c'_{-i} \in A'_{-i}$. Then $c'_{-i} = (c_{-i}, y_{-i})$ for a $c_{-i} \in A_{-i}$ and a $y_{-i} \in \{+, -\}^{\mathcal{N}-i}$. Consider $(x_i, y_{-i}) \in \{+, -\}^{\mathcal{N}}$.

If $(x_i, y_{-i})_{\mathcal{G}}$ is constant, then, by definition, it holds that $d'(a_i, x_i, c_{-i}, y_{-i}) = d(a_i, c_{-i})$ and $d'(b_i, x_i, c_{-i}, y_{-i}) = d(b_i, c_{-i})$. By supposition, it must be that $d(a_i, c_{-i}) \geq d(b_i, c_{-i})$. Hence, $d'(a_i, x_i, c_{-i}, y_{-i}) \geq d'(b_i, x_i, c_{-i}, y_{-i})$.

If $(x_i, y_{-i})_{\mathcal{G}}$ is not constant, then, by definition, it holds that $d'(a_i, x_i, c_{-i}, y_{-i}) = 0$ and $d'(b_i, x_i, c_{-i}, y_{-i}) = 0$. Hence, $d'(a_i, x_i, c_{-i}, y_{-i}) \geq d'(b_i, x_i, c_{-i}, y_{-i})$.

Hence, for all $c'_{-i} \in A'_{-i}$ it holds that $d'(a_i, x_i, c'_{-i}) \geq d'(b_i, x_i, c'_{-i})$. Therefore, $(a_i, x_i) \succeq_{M'_{\mathcal{G}}} (b_i, x_i)$.

(\Leftarrow) Suppose $(a_i, x_i) \succeq_{M'_{\mathcal{G}}} (b_i, x_i)$. Consider an arbitrary $c_{-i} \in A_{-i}$. Take a $y_{-i} \in \{+, -\}^{\mathcal{N}-i}$ such that $(x_i, y_{-i})_{\mathcal{G}}$ is constant. Then $(c_{-i}, y_{-i}) \in A'_{-i}$. By definition of d' , it holds that $d'(a_i, x_i, c_{-i}, y_{-i}) = d(a_i, c_{-i})$ and $d'(b_i, x_i, c_{-i}, y_{-i}) = d(b_i, c_{-i})$. By supposition, it must be that $d'(a_i, x_i, c_{-i}, y_{-i}) \geq d'(b_i, x_i, c_{-i}, y_{-i})$. Hence, $d(a_i, c_{-i}) \geq d(b_i, c_{-i})$. Hence, for all $c_{-i} \in A_{-i}$ it holds that $d(a_i, c_{-i}) \geq d(b_i, c_{-i})$. Therefore, $a_i \succeq_M b_i$.

(ii) Suppose $(a_i, x_i) \succ_{M'_G} (b_i, y_i)$. If $x_i = y_i$, then, by part (i) of this lemma, $a_i \succ_M b_i$.

Suppose $x_i \neq y_i$.

We first prove $a_i \succeq_M b_i$. Consider an arbitrary $c_{-i} \in A_{-i}$. There are two cases: (a) If $i \in \mathcal{G}$, then take a $z_{-i} \in \{+, -\}^{\mathcal{N}-i}$ such that $(y_i, z_{-i})_{\mathcal{G}}$ is constant. Then $(x_i, z_{-i})_{\mathcal{G}}$ is not constant. By definition of d' , it holds that $d'(a_i, x_i, c_{-i}, z_{-i}) = 0$. By supposition, $d'(b_i, y_i, c_{-i}, z_{-i}) = 0$. By definition of d' , it holds that $d(b_i, c_{-i}) = 0$. Hence, $d(a_i, c_{-i}) \geq d(b_i, c_{-i})$. (b) If $i \notin \mathcal{G}$, then take a $z_{-i} \in \{+, -\}^{\mathcal{N}-i}$ such that $(x_i, z_{-i})_{\mathcal{G}}$ is constant. Then $(y_i, z_{-i})_{\mathcal{G}}$ is also constant. By definition of d' , it holds that $d'(a_i, x_i, c_{-i}, z_{-i}) = d(a_i, c_{-i})$ and $d'(b_i, y_i, c_{-i}, z_{-i}) = d(b_i, c_{-i})$. By supposition, it holds that $d'(a_i, x_i, c_{-i}, z_{-i}) \geq d'(b_i, y_i, c_{-i}, z_{-i})$. Hence, $d(a_i, c_{-i}) \geq d(b_i, c_{-i})$. Hence, for all $c_{-i} \in A_i$ it holds that $d(a_i, c_{-i}) \geq d(b_i, c_{-i})$. Therefore, $a_i \succeq_M b_i$.

We now prove $b_i \not\prec_M a_i$. Because $(b_i, y_i) \not\prec_{M'_G} (a_i, x_i)$, there is a $c_{-i}^* \in A_{-i}$ and a $z_{-i}^* \in \{+, -\}^{\mathcal{N}-i}$ such that $d'(a_i, x_i, c_{-i}^*, z_{-i}^*) = 1$ and $d'(b_i, y_i, c_{-i}^*, z_{-i}^*) = 0$. Then $(x_i, z_{-i}^*)_{\mathcal{G}}$ is constant and $d(a_i, c_{-i}^*) = 1$. There are two cases: (a) If $i \in \mathcal{G}$, take a $z_{-i}^{**} \in \{+, -\}^{\mathcal{N}-i}$ such that $(y_i, z_{-i}^{**})_{\mathcal{G}}$ is constant. Then $(x_i, z_{-i}^{**})_{\mathcal{G}}$ is not constant. By definition of d' , it holds that $d'(a_i, x_i, c_{-i}^*, z_{-i}^{**}) = 0$. By supposition, $d'(b_i, y_i, c_{-i}^*, z_{-i}^{**}) = 0$. By definition of d' , it holds that $d(b_i, c_{-i}^*) = 0$. (b) If $i \notin \mathcal{G}$, then $(y_i, z_{-i}^*)_{\mathcal{G}}$ is constant. Because $d'(b_i, y_i, c_{-i}^*, z_{-i}^*) = 0$ and by definition of d' , it must be that $d(b_i, c_{-i}^*) = 0$. Either way, $d(b_i, c_{-i}^*) = 0$. Hence, there is a $c_{-i}^* \in A_{-i}$ such that $d(a_i, c_{-i}^*) = 1$ and $d(b_i, c_{-i}^*) = 0$. Therefore, $b_i \not\prec_M a_i$. \square

Theorem 2 Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model, let $\mathcal{G} \subseteq \mathcal{N}$, and let $M'_G = \langle \mathcal{N}, (A'_i), d', v' \rangle$ be the unit \mathcal{G} -transform of M . Let $R_1 \subseteq A \times A'$ be given by $\{(a, (a, x)) : a \in A \text{ and } x \in \{+, -\}^{\mathcal{N}}\}$. Then R_1 is an ι -bisimulation between M and M'_G .

Proof Suppose that $(a, a') \in R_1$ for $a \in A$ and $a' \in A'$. Then $a' = (a, x)$ for some $x \in \{+, -\}^{\mathcal{N}}$. Note that $a'_i = (a_i, x_i)$.

(i) Because $a' = (a, x)$, for all $p \in \mathfrak{P}$ it holds that $a \in v(p)$ iff $a' \in v'(p)$.

(ii) Suppose $a_i \notin \text{Adm}_M(i)$. Then there is a $b_i \in A_i$ such that $b_i \succ_M a_i$, that is, $b_i \succeq_M a_i$ and $a_i \not\prec_M b_i$. By Lemma 1(i), it holds that $(b_i, x_i) \succeq_{M'_G} (a_i, x_i)$ and $(a_i, x_i) \not\prec_{M'_G} (b_i, x_i)$, that is, $(b_i, x_i) \succ_{M'_G} (a_i, x_i)$. Hence, $(a_i, x_i) \notin \text{Adm}_{M'_G}(i)$ and therefore, $a'_i \notin \text{Adm}_{M'_G}(i)$.

Suppose $a'_i \notin \text{Adm}_{M'_G}(i)$. Then $(a_i, x_i) \notin \text{Adm}_{M'_G}(i)$. Then there is a $(b_i, y_i) \in A'_i$ such that $(b_i, y_i) \succ_{M'_G} (a_i, x_i)$. By Lemma 1(ii), we have $b_i \succ_M a_i$. Therefore, $a_i \notin \text{Adm}_M(i)$.

(iii) Consider an arbitrary $b \in A$. Let $b' = (b, x)$. Obviously, $b' \in A'$ and $(b, b') \in R_1$. Hence, there is a $b' \in A'$ such that $(b, b') \in R_1$.

(iv) Consider an arbitrary $b' \in A'$. It holds that $b' = (b, y)$ for some $b \in A$ and some $y \in \{+, -\}^{\mathcal{N}}$. Obviously, $(b, b') \in R_1$. Hence, there is a $b \in A$ such that $(b, b') \in R_1$.

(v) Suppose $a_i = b_i$ for an arbitrary $b \in A$. Let $b' = (b, x)$. Obviously, $b' \in A'$ and $(b, b') \in R_1$. Because $a_i = b_i$ and $x_i = x_i$, it holds that $a'_i = b'_i$. Hence, there is a $b' \in A'$ such that $a'_i = b'_i$ and $(b, b') \in R_1$.

(vi) Suppose $a'_i = b'_i$ for an arbitrary $b' \in A'$. It holds that $b' = (b, y)$ for some $b \in A$ and some $y \in \{+, -\}^{\mathcal{N}}$. Obviously, $(b, b') \in R_1$. Because $a'_i = (a_i, x_i)$ and $b'_i = (b_i, y_i)$, it must be that $a_i = b_i$ and $x_i = y_i$. Hence, there is a $b \in A$ such that $a_i = b_i$ and $(b, b') \in R_1$.

Therefore, $(M, a) \rightleftharpoons_i (M''_{\mathcal{G}}, (a, x))$. \square

Lemma 2 Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model, let $\mathcal{G} \subseteq \mathcal{N}$, and let $M''_{\mathcal{G}} = \langle \mathcal{N}, (A'_i), d'', v'' \rangle$ be the zero \mathcal{G} -transform of M . Let $a_i, b_i \in A_i$ and $x_i, y_i \in \{+, -\}$. Then

- (i) $a_i \succeq_M b_i$ iff $(a_i, x_i) \succeq_{M''_{\mathcal{G}}} (b_i, x_i)$
- (ii) If $(a_i, x_i) \succ_{M''_{\mathcal{G}}} (b_i, y_i)$, then $a_i \succ_M b_i$.

Proof Analogous to the proof of Lemma 1. \square

Theorem 3 Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ be a deontic game model, let $\mathcal{G} \subseteq \mathcal{N}$, and let $M''_{\mathcal{G}} = \langle \mathcal{N}, (A'_i), d'', v'' \rangle$ be the zero \mathcal{G} -transform of M . Let $R_2 \subseteq A \times A''$ be given by $\{(a, (a, x)) : a \in A \text{ and } x \in \{+, -\}^{\mathcal{N}}\}$. Then R_2 is an ι -bisimulation between M and $M''_{\mathcal{G}}$.

Proof Analogous to the proof of Theorem 2. \square

References

- Arrow, K. J. (1951). Alternative approaches to the theory of choice in risk-taking situations. *Econometrica*, 19, 404–437.
- 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: Teams and frames in game theory*. Princeton University Press.
- Belnap, N., Perloff, M., & Xu, M. (2001). *Facing the future: Agents and choices in our indeterminist world*. Oxford University Press.
- Bhargava, R. (1992). *Individualism in social science: Forms and limits of a methodology*. Clarendon Press.
- Blackburn, P., de Rijke, M., & Venema, Y. (2001). *Modal logic*. Cambridge University Press.
- Bratman, M. E. (2014). *Shared agency: A planning theory of acting together*. Oxford University Press.
- Duijf, H. (2018). *Let's do it! Collective responsibility, joint action, and participation*. Ph.D. thesis, Utrecht University.
- Elster, J. (1985). *Making sense of Marx*. Cambridge University Press.
- Elster, J. (1989). *Nuts and bolts for the social sciences*. Cambridge University Press.
- Horty, J. F. (1996). Agency and obligation. *Synthese*, 108, 269–307.
- Horty, J. F. (2001). *Agency and deontic logic*. Oxford University Press.
- Horty, J. F., & Belnap, N. (1995). The deliberative stit: A study of action, omission, ability, and obligation. *Journal of Philosophical Logic*, 24, 583–644.

- Kanger, S. (1957). New foundations for ethical theory. Privately distributed pamphlet. In R. Hilpinen (Ed.), *Deontic logic* (pp. 36–58). D. Reidel Publishing Company.
- Kincaid, J. (1986). Reduction, explanation, and individualism. *Philosophy of Science*, 53, 492–513.
- Kohlberg, E., & Mertens, J.-F. (1986). On the strategic stability of equilibria. *Econometrica*, 54, 1003–1037.
- Kooi, B., & Tamminga, A. (2008). Moral conflicts between groups of agents. *Journal of Philosophical Logic*, 37, 1–21.
- List, C., & Pettit, P. (2011). *Group agency: The possibility, design, and status of corporate agents*. Oxford University Press.
- List, C., & Spiekermann, K. (2013). Methodological individualism and holism in political science: A reconciliation. *American Political Science Review*, 107, 629–643.
- Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. Wiley.
- McNamara, P. (1996). Must i do what i ought? (Or will the least I can do do?). In M. A. Brown & J. Carmo (Eds.), *Deontic logic, agency and normative systems* (pp. 154–173). Springer.
- Nagel, E. (1961). *The structure of science. Problems in the logic of scientific explanation*. Harcourt, Brace & World, Inc.
- Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. The MIT Press.
- Pörn, I. (1970). *The logic of power*. Basil Blackwell.
- Quinton, A. (1975). Social objects. *Proceedings of the Aristotelian Society*, 76, 1–27.
- Savage, L. J. (1972). *The foundations of statistics* (second revised). Dover Publications, Inc.
- Sawyer, K. R. (2002). Nonreductive individualism: Part I—Supervenience and wild disjunction. *Philosophy of the Social Sciences*, 32, 537–559.
- Sawyer, K. R. (2003). Nonreductive individualism: Part II—Social causation. *Philosophy of the Social Sciences*, 33, 203–224.
- Schelling, T. C. (1960). *The strategy of conflict*. Harvard University Press.
- 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.
- Tamminga, A., Duijf, H., & Van De Putte, F. (forthcoming). Expressivity results for deontic logics of collective agency. *Synthese*. <https://doi.org/10.1007/s11229-020-02597-0>
- Tamminga, A., & Hindriks, F. (2020). The irreducibility of collective obligations. *Philosophical Studies*, 177, 1085–1109.
- Tuomela, R. (2013). *Social ontology: Collective intentionality and group agents*. Oxford University Press.
- Udehn, L. (2001). *Methodological individualism: Background, history, and meaning*. Routledge.
- van Benthem, J. F. A. K. (1984). Correspondence theory. In D. Gabbay & F. Guenther (Eds.), *Handbook of philosophical logic* (Vol. II, pp. 167–247). D. Reidel Publishing Company.
- Van De Putte, F., Tamminga, A., & Duijf, H. (2017). Doing without nature. In A. Baltag, J. Seligman, & T. Yamada (Eds.), *Logic, rationality, and interaction: 6th international workshop* (pp. 209–223). Springer.
- van Kutschera, F. (1986). Bewirken. *Erkenntnis*, 24, 253–281.