

Article

A Relational Perspective on Collective Agency

Yiyan Wang ^{1,2,*}  and Martin Stokhof ^{1,2,†} 

¹ Tsinghua-Amsterdam Joint Research Centre for Logic, Department of Philosophy, Tsinghua University, Beijing 100084, China

² Institute for Logic Language and Computation, Department of Philosophy, University of Amsterdam, 1012 WX Amsterdam, The Netherlands; m.j.b.stokhof@uva.nl

* Correspondence: wang-yy19@mails.tsinghua.edu.cn

† These authors contributed equally to this work.

Abstract: The discussion of collective agency involves the reduction problem of the concept of a collective. Individualism and Cartesian internalism have long restricted orthodox theories and made them face the tension between an irreducible concept of a collective and ontological reductionism. Heterodox theories as functionalism and interpretationism reinterpret the concept of agency and accept it as realized on the level of a collective. In order to adequately explain social phenomena that have relations as their essence, in this paper we propose a relational, holistic account of collective agency and argue that functionalism and interpretationism can be integrated into such an account.

Keywords: collective agency; collective intentionality; holism; interpretationism; functionalism

1. Introduction

In the broadest sense, the term “agent” refers to those who have the ability to act. Human agency further requires the capacity of rationality and its manifestation. As the name suggests, collective agency means realizing this capacity at the collective level. The interpretation of agency involves a core concept in philosophy of action: intentionality.¹ Similarly, at the collective level, the interpretation of collective agency involves a parallel concept: collective intentionality.

Research on collective intentionality has flourished in recent decades. Compared with the traditional social science that insists on individualism (see Popper [1]), research on collective intentionality reflects a strong tendency towards non-individualism. The work of Michael Bratman, Margaret Gilbert, Raimo Tuomela, John Searle and others has considerably broadened our horizons in social philosophy and made people generally accept that intentionality can be shared in social actions. However, in these approaches, the influence of individualism still lingers and that leads them to regard individuals as the cornerstone of the entire theoretical edifice and bearers of collective intentions. And with that, they encounter difficulties facing the tension between an irreducible concept of collective intentionality and ontological individualism.

Recently, some alters have modified the definition of agency and relaxed certain restrictions, for example that any intentional subject should meet the same physical or psychological standards as individuals do, thereby making the existence of collective agency possible. Two notable forms of realism with respect to collective agency are interpretationism (cf. Rovane [2]; Tollefsen [3]) and functionalism (cf. List and Pettit [4]). From the perspective of interpretationism, as long as a system behaves in a way that can be interpreted as rational by others, it can be regarded as an agent. Interpretationism focuses on organizational structures of collectives and believes that these structures guarantee the possibility of rational perspectives on collectives. On the other hand, functionalism believes that agency is embodied in the agent’s ability to perform actions that meet certain functional conditions, such as representation, motivation, and rational capacity. Analogously,



Citation: Wang, Y.; Stokhof, M. A Relational Perspective on Collective Agency. *Philosophies* **2022**, *7*, 63. <https://doi.org/10.3390/philosophies7030063>

Academic Editors: Joseph Ulatowski and Marcin J. Schroeder

Received: 29 March 2022

Accepted: 6 June 2022

Published: 8 June 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

functionalism characterizes collective agency in terms of its action results and explains collective agency in terms of reliable aggregation of converging opinions of the members of a collective to ensure that a collective's opinions are rational.² In this paper, it is argued that these two different perspectives on realism of collective agency can be integrated into a relational account.

Various authors (cf. Baier [5]; Meijers [6]; Schmid [7]) have claimed that we need a relational or holistic explanation of collective intentionality. In line with their arguments, we will demonstrate that collective agency essentially is a combination of relational eventualities³ of a collective consisting of members of a collective, the targets of collective actions, structures and functions of a collective. Moreover, for each collective agent, these relational eventualities constitute a relational identity; not a physical existence, such as a table, but a non-physical existence such as friendship, policy or a nation-state. Basically, our claim is that this relational identity cannot be reduced to a simple aggregation of its members, plus the targets of collective actions, plus structures, plus functions.

Before starting the discussion, two premises need to be stated: 1. we only focus on collectives containing specific institutional structures or functions; in other words, a simple collection of individuals does not make a collective; 2. our discussion does not involve spontaneous behaviors, which, in our opinion, constitute an aspect of individual intentionality, which concerns the accessibility and interrelation of the mental and material worlds. However, in this paper, we focus on the issue of collective intentionality that is on the relation between individual minds and social phenomena.

The structure of the paper is as follows. In the next section, we will review orthodox theories of collective intentionality and explain why they cannot avoid what is called the "circularity problem". Furthermore, we will discuss why most philosophers are resistant to the existence of collectives in Section 3. Specifically, we will start with an objection to the basic argument of physicalistic individualism by showing that the very same argument does not exclude the existence of collective. In Section 4, we will return to the existing forms of realism with respect to collective agency and integrate them into a relational account. In Section 5, we will respond to the objection that the existence of collective agency will harm individual agency and discuss how collective agency affects individual actions practically in a relational view. A summary and possible future directions are included in the conclusion.

2. Irreducible Concepts of Collectives

As Petersson [9], Schmid and Schweikard [10] have pointed out, orthodox theories of Gilbert, Tuomela, Searle and Bratman are trapped in the circularity problem because they persist on individualism in varying degrees. At least two types of individualism need to be distinguished. Ontological individualism holds that a collective is nothing more than a complex of individuals and their actions. Methodological individualism insists that social phenomena can only be explained by showing how they arise from individual actions and the intentional states that underlie and motivate those actions.

Generally speaking, all these theories adhere to ontological individualism. Among these, Gilbert, Tuomela and Searle admit the existence of an irreducible concept of a collective in a methodological sense to some extent. In Gilbert, it is "joint commitment"; in Tuomela, it is "we-intention".⁴ Searle further binds the concept of "we-intention" to individuals' physical brains or even brains in vats, manifesting an apparent tendency of internalism. Bratman goes even further by insisting that we can explain collective intentionality without any irreducible concept of a collective; that is, any collective intentionality can be explained by its members' individual intentionality and interdependent relationships. Using these characteristics, we will elucidate how they fall prey to the dilemma of circularity.

2.1. Ontological Individualism

Margaret Gilbert holds that activities of a collective require the collective to take on a "joint commitment", which in her works means "a kind of commitment of the will. In this

case, the wills of two or more people create it, and two or more people are committed by it" ([13], p. 134). Such a joint commitment implies a normative relationship among members that each has an obligation to act accordingly and demands others' conforming actions. She further defines "plural subject" as people "are jointly committed to doing something as a body" ([13], p. 145). Although Gilbert claims we need to transcend individualism (cf. [14]), she restricts the conceptual basis of "joint commitment" to include only conditional personal commitments (cf. [15]) and explicitly bases her analysis on "the concept of an individual person with his own goals, and so on, does not require for its analysis a concept of a collectivity" ([16], p. 435).

Among various criticisms of Gilbert's theory, there is a particular criticism that claims that Gilbert faces the danger of circularity. It asks how a joint commitment to jointly perform an action can be formed without presupposing the concept of a joint action that is meant to be brought about by the plural subject; for, if we say individual agents *A* and *B* jointly commit to have a meeting, the concept of the joint action of their meeting is presupposed. It seems that joint commitment presupposes rather than explains the concept of a joint action it is intended to analyze. A possible reply could be that before *A* and *B* have the joint commitment and act as a body, each of them is already willing to do so. However, such a reply is not enough because such circularity reappears at the level of willingness. Nevertheless, her theory of creating joint commitments is a normative description rather than a conceptual analysis (cf. [9]). Thus, it would not be proper to criticize her theory with circularity problems.

Regarding Tuomela's theory, it distinguishes the form of individual attitudes in "we-mode" versus "I-mode", as "notions of having an attitude and acting as a group member versus as a private person" ([17], p. 46). Moreover, he defines "we-intention" as follows: an individual agent has the we-intention to do *X* if and only if they have a respective individual intention to do *X* and have certain beliefs about the actions and beliefs of other participants, and there is a mutual belief among the participants about this. "When there is shared we-intention in this sense we can speak of at least a weak joint intention, one in the I-mode. In the we-mode case we can say that the participant intends to participate in *X* at least in part because the group intends to perform *X*" ([18], pp. 41–42). Tuomela's analysis of (we-mode) "we-intention" can be briefly characterized as follows: A member *A* of a group *G* has we-intention to do action *X* if and only if (1) *A* intends to do their part of *X*;⁵ (2) *A* believes that others in *G* will do their parts; (3) *A* believes that there is a mutual belief among the participating members of *G*; (4) (1) in part because of (2) and (3) ([18], pp. 42–43; [19], pp. 340–341).

Observe that although Tuomela clarifies that "my analysis is not meant to be reductive but is rather meant to elucidate the irreducible notion of we-intention in a functionally informative way" ([19], p. 358), he also explicitly mentions that "the present approach is ontologically individualistic in the sense that we-mode states and properties are attributed to individuals, severally or jointly, when they function as group members" ([20], p. 93). In other words, he believes that although we have a special kind of attitude—we-mode—and a special kind of intention—we-intention—this does not presuppose any existence of collective entities as its bearer.

Tuomela's theory also faces a problem of circularity. By his analysis, if an agent *A* intends to do their part of *X*, and has we-intentions that satisfy the conditions (1)–(4) above, it still presupposes a concept of a collective that intended to be constituted by we-intention. So, "we-mode" and "we-intention" do not constitute the concept of a collective, but rather presuppose its existence. In ([19], pp. 355–361), he replies that his account does not contain any vicious circularity since the analysis means that a collective is consisted in, instead of constituted by, the participants' having we-intention. However, this is not satisfactory since it is just a verbal way out; he still needs to explain what "consisted in" is.

2.2. *Ontological Individualism with Internalism*

Searle supports the idea of “we-intention” and rejects Tuomela’s analysis of it; he says, “this account is typical in that it attempts to reduce collective intentions to individual intentions plus beliefs. I, on the contrary, am proposing that no such reduction will work, that ‘we-intentions’ are primitive” ([21], p. 404). However, meanwhile Searle insists that “all intentionality, whether collective or individual, has to exist inside individuals’ head” ([22], p. 44). To Searle, the primitive, irreducible “we”-form is the core feature of cooperation, yet this concept of a collective can only exist in the participants’ heads and motivate their contributions. Although insisting on an irreducible concept of a collective, he claims that his theory adheres to methodological individualism (cf. [21]). In addition, the above conception of individuals as just brains has a strong flavor of internalism.

The combination of internalism and ontological individualism makes his theory more vulnerable to criticism. In addition to the circularity problem, Searle also has to accomplish a seemingly impossible thing: to explain collective actions, which are characterized as interpersonal, in an internalistic way. Moreover, it remains an open question how an irreducible, primitive “we”-form emerges in the individual brain.

2.3. *Ontological Individualism with Methodological Individualism*

Bratman executes a thorough reductionistic analysis in which a collective intention can ultimately be reduced to interrelated individual intentions (cf. [23]). He suggests that each participant agent can intend not only their respective contributions but also a joint activity *J* as such. His analysis runs as follows: “We intend to *J* if and only if 1. (a) I intend that we *J* and (b) you intend that we *J*; 2. I intend that we *J* in accordance with and because of 1a, 1b, and meshing subplans of 1a and 1b; you intend that we *J* in accordance with and because of 1a, 1b, and meshing subplans of 1a and 1b. 3. 1 and 2 are common knowledge between us” ([24], p. 121). Although Bratman claims his theory can radically reduce collective intentionality to the individual level, many situations with a concept of a collective do not seem to yield to his reduction.⁶ In addition, the collective concept “we *J*” is not fully reduced in his analysis: it reappears at the individual level.

The criticism of circularity is brought against Bratman’s work and takes the form of a question: how individuals can refer to joint activity in the absence of “jointness”? How can an agent assume that others will intend as they do without assuming there is a collective intention? It seems individual intentions, as he conceives of them, may not be a sound ground for shared intention but rather a superfluous expression of existing collective intentionality. Bratman replies that in order to achieve “jointness”, members’ intentions also need to satisfy the following conditions: “each is appropriately responsive to the other; each is set not to thwart the other and is at least minimally disposed to help the other if needed; there is (feasibility-based) persistence interdependence in relevant intention; there is, as a result of all this, rational pressure in the direction of social rationality; each expects all of this to issue in the joint action each intends; and all this is out in the open” ([25], p. 95). These conditions are, in his terminology, a form of interconnected functioning between the intentions and actions of the participants. However, again, this interconnected process of intention formation presupposes the concept of a collective it is meant to constitute.

2.4. *The Gordian Knot*

As mentioned above, centering on ‘joint activity’, ‘we-mode’, or ‘joint commitment’ while retaining the ‘I’-form of the constitute individual intentions will face the danger of circularity. However, this is not the crucial difficulty of individualism, in our opinion.

Humberstone [26] distinguishes analytical circularity from inferential circularity and claims that the latter is generally vicious, but the former is not. Suppose the general form of an account of a concept *K* runs as follows: the concept *K* applies if and only if certain conditions C_1, C_2, \dots, C_n are obtained. Such an account can be thought of as a putative analysis of *K*. Viewed as such, an account is analytical circular if the concept *K* is employed in specifying the conditions C_1, C_2, \dots, C_n . Alternatively, we can also think

of such an account from the perspective of the bi-directional claim “if and only if.” The direction of “only if” just provides necessary conditions for applying the concept *K*. When it comes to the direction of “if”, the situation is different in a crucial sense. As Humberstone formulates it: “the kind of circularity pertinent here will not be like the circularity of a putative analysis or definition, but rather, like the circularity of an argument” ([26], p. 250). That conditions C_1, C_2, \dots, C_n obtain are premises of this argument, and that concept *K* applies is its conclusion. Thus “an account of the application some concept is inferentially circular when any argument or inference from premises claiming the various conditions provided by the account to obtain, to the conclusion that the concept applies, is itself circular in whatever sense an inference or argument can be circular” ([26], p. 250). A more usual way to explain the flaw of circularity is this: if a concept is being explained, the explanation should not be understood only by those already possessing the concept. In this sense, the analytical circularity is deficient in the fact that it impedes the transition from an understanding of the terminology in an account to an understanding of concepts. In contrast, the flaw of the inferential circularity is not only at the level of understanding but also at the level of knowledge: people cannot obtain any new knowledge from a definition that is inferentially circular.

Under such a distinction, Petersson [9] mentions that the above analyses of acting with collective intentions face analytical circularity. For example, if we say one of the necessary and sufficient conditions for “I push the car intentionally” to be true is that I know I push the car intentionally; that is the case of inferential circularity. Because if I know I push the car intentionally, I already push the car intentionally. Such an analysis may not contribute to our understanding of the term “intentional”. However, if we say one of the necessary and sufficient conditions of “we push the car collectively” to be true is that our pushing the car is explained by “joint commitment”, “we-intention” or “we-*J*” and further involves each individual’s intending on pushing the car collectively, although the concept analyzed is implicitly present in components of the analysis, we do obtain further knowledge compared to the single statement of “we push the car collectively”.

In line with Humberstone and Petersson’s arguments, we also think the so-called circularity problem these orthodox theories encounter is not severe, since they (at most) presuppose the existence of a concept of a collective instead of the actual existence of a collective. After all, concepts can exist without any assumption that there is something ontological that corresponds to them. Thus the object being explained is different from the one being assumed conceptually, and we do obtain new knowledge.

Furthermore, we believe that the true Gordian knot lies in the attempt to explain collective agency or intentionality solely in terms of individual agency or intentionality. In such an attempt, an irreducible concept of collective agency remains at the individual level, and then the question of where that concept comes from arises. If there is, in fact, no such thing, as ontological individualism claims, the concept of a collective that individuals have should be considered as nothing but an illusion. However, we still need to explain why that illusion exists. The crucial point is why this illusion of a collective plays an essential role in collective actions. The three available answers to the question of the essential role of the concept of a collective are non-satisfactory; they are (1) claim that they are mere illusions; (2) acknowledge that the concept of a collective that cannot be reduced; (3) reduce the concept of a collective to the individual level to dismiss it. The “non-satisfactory” here means that they may not contribute to our understanding of the essential role of the concept of a collective. If we admit that answer (1) is irrelevant to our understanding and that the reduction of approach (3) cannot be completed (as in Bratman’s theory, “we *J*” reappears at the individual level), only (2) is left to be elaborated. Suppose we put “Acknowledge the concept of a collective that cannot be reduced” as one of the necessary and sufficient conditions for, “The concept of a collective plays an essential role in collective actions”. Then it involves inferential circularity rather than merely analytical circularity. This is because, by referring to concepts with irreducibly collective content, we already presuppose that some kind of conception of collective action plays an essential role without

providing further analysis of it. In other words, we cannot obtain any new knowledge from answer (2). The whole individualistic explanation fails to provide an understanding of the concept of a collective and its source. This is the crucial challenge for individualistic accounts, especially for those that admit an irreducible concept of a collective. The relevant circularity is not an analytical one but an inferential one. To cut this Gordian knot, we need to explore its source related to our ontological assumptions, which will be discussed in the next section.

3. Why Can't We Speak of Collective Entities?

Before presenting our point of view in Section 4, we first address the question of why individualism has such a broad audience. Baier [5] points out that a Cartesian specter has haunted the field of collective agency for a long time and has led to the existing individualistic theories. Meijer [6] and Schmid [7] continue to develop this view and refute any individualistic tendencies in the discussion of collective agency. We believe that resisting this Cartesian influence will shed new light on our breaking away from the shackles of the physical bearer to make it possible to acknowledge the existence of collectives. We will illuminate this by reformulating the basic argument of physicalistic individualism and showing that it cannot exclude the existence of collectives.

3.1. *The Basic Argument of Physicalistic Individualism*

People always regard intentionality as a mental state possessed by an individual. Correspondingly, we intuitively find it is challenging to accept collective intentionality if this indicates a certain spiritual body of a collective: a Hegelian spirit of a collective possesses such collective intentionality, just as an individual possesses their own intentionality. The Cartesian specter plays a decisive role in this. In *Meditations*, Descartes' argument of "Cogito ergo sum" constructs the concept of an independent individual mind, which is isolated from the social and natural world. In his narrative, intentionality is not a property of interpersonal and multi-subjective types of entities. As formulated by Schmid, this tradition brings about a kind of "formal individualism", which restricts intentionality to an 'I'-form and rejects any plural form of intentionality.

Bratman's reductionist theory presents firm ontological individualism plus methodological individualism. Searle's strategy of looking for brains in vats as bearers is obviously internalistic. Gilbert, Tuomela and even Sugden, Gold, etc., (in the economics tradition of team reasoning) adopt a modest approach: they accept concepts of collectives but still insist on ontological individualism. The key observation is that these accounts avoid attributing any concept of a collective (whether reducible or not) to a non-individual collective entity. The reason behind this is that all the authors accept that nothing at the collective level can serve as a vehicle for concepts of a collective. Per contrast, the individual in a physical sense is a perfect vehicle, so in addition to individual intentionality, they ascribe part or all of collective intentionality to the individual, or more directly, to the individual brain. Note that there is a strong physicalist or materialist streak behind each of these theories, even for those who take a modest approach; without loss of generality, we may name such a type of theory more properly as "physicalistic individualism".

Both Cartesian internalism and contemporary physicalism lurk behind individualistic accounts. For narrative convenience, here we briefly reconstruct the basic argument of contemporary physicalistic individualism,⁷ which consists of an internalistic argument of others' having intentionality and a reinforced physicalistic condition. The former contains three premises (P1, P2, P3) and a conclusion (C1):

- (P1) Internalism in principle: the justifications of our cognitive contents are independent of the outside world.
- (P2) Self-awareness: everyone is familiar with themselves and can be aware of their physical body and their cognitive contents, and understand the correlations and differences between them.

(P3) Awareness of others: we can see and touch other people (in a physical sense), although we cannot see and touch cognitive contents of other people, we can be aware of such things. And we can also have the concept of similarity and difference with other humans.

(C1) With (P2) and (P3) as basic intuitive facts, each individual “I” knows themselves well and is aware of others. This “I” would see others as familiar kinds of stuff and attribute similar cognitive contents to them. According to (P1), “I” do not need any external source to justify knowledge (internalism). Therefore, every individual has the concept of other people having intentionality, independent from any actual connection with such intentionality.

Based on the facts of (P2) and (P3), Cartesian internalism acknowledges the awareness of others having intentionality.

The conclusion (C1) easily fits with physicalism and further strengthens it to a claim of physicalistic individualism. The reinforced physicalistic condition is as follows (P4):

(P4) Physicalistic function of brains: all intentionality requires a physical brain, and in the end is a physical existence.

(C2) With (C1) as the basic conclusion, according to (P4), such intentionality depends only on physical brains, which can only own by individuals. In other words, intentionality only exists on the individual level (me and other persons, even intelligent animals), which is a typical individualistic claim.

There are at least three aspects of this argument of physicalistic individualism that can be thought of differently. Of course, the first is directed at the condition (P1) and is adopted by the authors that oppose the Cartesian tradition mentioned at the beginning of this section. In this way of thinking, theories of collective agency do not succumb to those individualistic constraints and are not limited to dealing with the tension between irreducible concepts of collectives and ontological reductionism. It allows us to embrace a more general relational, holistic account that can explain collectives through primitive relational concepts. The second alternative view is connected with (P4) and argues that a system could have intentionality or agency without any physical bearers and that collective agency can exist in the sense of interpretationism or functionalism.⁸ We will come back to these accounts in detail in Section 4.

Nevertheless, we think these two alternative ways of thinking are not sufficient to turn our attention to the relational account because they are alternatives to, instead of refutations of, physicalistic individualism itself. Even if we claim that individualism will face intractable theoretical difficulties and appeal to relationalism or holism to explain collectiveness, individualists can still find a way to escape, such as relaxing restrictions of methodological individualism (as in the work of Gilbert and Tuomela) and claiming that the relational account is nothing but another narrative.

3.2. Objection to Familiarity

In our opinion, the real problem with physicalistic individualism is that the logical pattern behind it is invalid. That pattern is: (1) “I” am familiar with myself as an agent; (2) as an agent “I” have intentionality; (3) in the same way as with myself, “I” am familiar with others as agents; (4) thus other agents also have intentionality; (5) and those agents that “I” am not familiar with do not have intentionality.

The problem here is that familiarity requires a premise that the objects being compared are of the same type. At first glance, it certainly seems that “I” and other people belong to the same type of human being. However, knowledge of these two is of different types. Our knowledge of ourselves is intuitively distinct from our knowledge of external objects.⁹ Instead of the actual existences of “I” and others, what the internalist agent compares are two different types of knowledge of those existences. Thus, if this distinction holds, the extension from first-person to third-person having intentionality is problematic;¹⁰ in other words, familiarity does not apply here.

In the argument of physicalistic individualism, the sentence in (C1) “This “I” would see others as familiar kinds of stuff and attribute similar cognitive contents to them” follows the logical pattern of familiarity. The result of mistakenly applying familiarity to two different types of knowledge is to accept a strict definition of an agent, i.e., all agents must have the exact attributes and organizational structures as the canonical agent such as myself, including necessary physical bearers. Hence, physical brains become a necessary condition for collective agency. Nevertheless, this is absurd. Our awareness of collectives and other people’s intentional states is of the same invisible and untouchable objects. Yet, just because “I” am familiar with a particular relationship between my physical brain and my cognitive contents, “I” takes for granted that the physical brains of others also have their cognitive contents. However, since a collective does not have a physical brain, although “I” can be aware of a so-called collective, “I” firmly believes it does not exist and cannot have intentionality because “I” am not familiar enough with it compared to a canonical agent like myself.

Based on the distinction of two types of knowledge, we do not need to stick to the definition of an agent in the strict sense. This means our understanding of an agent can be loosened to a level that an agent is only assumed to have sufficiently similar properties and functions as canonical agents; then, we do not need to seek any physical brains for collective agents.

Now the severe problem is that if we abandon the logical pattern of familiarity, we can declare collective existence through a similar argument without disposing of the premise of Cartesian internalism:

- (P1) Internalism in principle: the justifications of our cognitive contents are independent of the outside world.
- (P2) Self-awareness: everyone is familiar with themselves and can be aware of their physical body and their cognitive contents, and understand the correlation and difference between them.
- (P3*) Awareness of collectives: although we cannot see collectives (physically), any person-in-society does perceive the existence of collectives, at least in the conceptual sense. Moreover, a person-in-society can be aware of a different kind of intentionality than that in a collective sense.¹¹
- (C*) With (P2) and (P3*) as basic intuitive facts, each individual agent “I” knows myself well and is aware of collectives. Since the definition of an agent no longer requires physical grounds, “I” have no reason to reject the existence of collectives, at least in the conceptual sense. In this way, “I” will realize that such concepts arise from the manifestation of some counterpart of the physical world different from individuals, to which “I” would naturally attribute collective intentionality. According to (P1), justifications of my knowledge are independent of the outside. Therefore, collectives may exist and have their intentionality, and such intentionality depends on something different from individuals; in other words, intentionality can be realized at the level of a collective.

It is an apparent problem that internalism will make our cognition completely independent of any relation that refers to the external. In fact, concerning existence in the ontological sense, internalists cannot obtain any practical knowledge except of themselves. Therefore, the argument for others having intentionality also applies to a collective. In addition, physicalists think that they can find physical grounds for others’ intentionality. Although we cannot see or touch collectives and thus cannot find a solid counterpart of collective intentionality, by internalism, our awareness of the existence of concepts of collectives would suggest that we ascribe collective intentionality to the ontological source of such concepts. Suppose there is a so-called dependency between the cognitive contents and the physical body of each individual (and such dependency is essentially an assumption of analogy to myself and may not be confirmed at all). In that case, collective intentionality can also depend on the source of concepts of collectives instead of physical brains.

If the above argument is valid, we may conclude that physicalistic individualism does not bring any substantive constraint. With the premise of internalism and a correct understanding of familiarity, we can conclude that collective intentionality can be attributed to the non-individual source of irreducible concepts of collectives. And that conflicts with individualism. It needs to be pointed out that such an objection against familiarity does not necessarily lead to collective existence. Philosophers who are worried about the collective specter can rest assured. It can only adequately show the absurdity of physicalistic individualism. Adhering to it can neither solve the problem of circularity nor naturally explain collectiveness. This suggests that it may be worthwhile to investigate whether a non-individualistic, relational and holistic account that is more suitable for explaining social phenomena should be adopted. Under such a relational account, we believe that collective intentionality depends on the collection of all the relational eventualities that generate it.

4. The Relational Account of Collective Agency

Orthodox theories that adhere to individualism are based on an intuitive argument about intentionality: whether a subject has an intention is based on whether it has a mind. However, two heterodox theories that acknowledge the existence of collective agency and collective intentionality emerged: functionalism and interpretationism.

4.1. Functionalism and Interpretationism

The classic functionalist approach of agency believes that a system manifesting agency only requires its behaviors to instantiate the agential model, which is represented as its behaviors meeting some necessary functional conditions. According to contemporary functionalist theory, the mind is an entity that can operate through its functional organization. In a holistic view, each part of the entity is defined by its role in the entire system, and these roles can be multiply realized.¹² Specifically, Pettit and List believe that a functionalist account of agency should abandon physical constraints: agents must be composed of certain material substance that is the same as that of canonical agents such as humans or animals. Additionally, they believe we should also abandon psychological mode constraints: the process by which the agent forms its inclination to act must be realized through a certain psychological model, such as consciousness. Moreover, a functionalist account should turn to the only condition of ratiocinative capacity: as long as the subject's actions reflect its capacity of rationality, it can be identified as an agent. "The system has to be able to identify some of the demands imposed by the pattern¹³ as regulative or normative requirements, and to let the identification of those demands reinforce conformity and underpin the recognition of non-conformity as a failure" [32]. Therefore, their theory relaxes the restrictions on agency and maintains that as long as collective actions reflect the collective's rational ability, the collective can have intentional states and agency. On this basis, their theory examines the issue of how to achieve rationality in collective decision making in voting and claims that if multiple decisions of a collective are logically consistent,¹⁴ they reflect rationality and, by their definition, can be treated as a collective agent (cf. Pettit [33]; List and Pettit [4]).

On the other side, interpretationism maintains that if we can interpret an agent, then the agent is intentional,¹⁵ and it asks: what assumptions do we need to explain agential behaviors? If we assume that certain premises and successfully explain agential behaviors in terms of them, then our assumptions about agents can be justified, thereby determining the nature of agents. The common assumption is that an agent should obey rational norms, i.e., is rational. Such an assumption is basic since we cannot understand and interpret any agent until we make it. Contemporary interpretationism follows a view called Neo-Lockeism: personhood means to keep a rational first-person point of view. They believe that this first-person perspective is not related to the unified consciousness that Locke originally insisted on, nor is it related to souls or brains, but is only a perspective that can be applied and understood by others. This rational perspective does not need to be

bound to any individual soul or animal (cf. Rovane [2]). Based on the assumption of a rational perspective, interpretationism attributes beliefs, desires, and intentions to agents according to their environment and functions and can further predict agents' behaviors based on those attributed intentions. Furthermore, "these attributions are not made in isolation but holistically" [3]. Because there is no link between physical and spiritual aspect, contemporary interpretationism recognizes the existence of collective agency, in the sense of realizing the rational first-person point of view at the collective level. In its analysis of collective intentions, the structure of a collective is regarded as an essential component. "The structure of the organization provides a way of synthesizing the disparate perspectives of individuals into a unified perspective from which goals and subgoals can be set and achieved" [3]. Individuals can think and act intentionally in the sense of "collectively" by adopting the collective perspective. Interpretationism believes that a collective is rational, if and only if, it has a certain structure that enables it to achieve its goals effectively. When this hypothesis is established, interpretationism "attributes beliefs, intentions, and desires to organizations the same way we do to individuals" [3].

Interpretationism is a mild realism of agency. It uses a hypothesis-inference approach to judge what assumptions are necessary in a roundabout way. In interpretationism, we can directly describe the attribution of beliefs, desires, intentions to collective agents. On the other hand, functionalism is more a kind of behaviorist judgment about agency. It judges whether a collective has agency by determining whether its actions meet necessary functional conditions. Because functionalism relies only on observable actions, it can be well integrated with specific approaches in the social sciences such as game theory and social choice theory.

4.2. Turn to a Unified Account

In this section we propose a new perspective that starts from the observation that interpretationism and functionalism are different but compatible and can be integrated into a unified account. In our opinion, the most significant difference between them is in their research perspective. In addition to relaxing the physical and psychological constraints on agency, contemporary theories of interpretationism and functionalism loosen the constraints on agency from two aspects: interpretationism relaxes the constraints on the conditions that behaviors of a collective should satisfy; functionalism relaxes the constraints on the internal structure of a collective. If our understanding is correct, it is possible to integrate them into one unified account in which the essence of collective agency is directly revealed.

Our account is based on the observation that we all have no problem accepting and using concepts of collectives in our daily lives, for instance, companies, institutes, nation-states, etc., and also that we have the ability to attribute collective beliefs, desires, etc., to collective agents, as well as to predict and expect their behaviors. For example, "The soccer team has the determination to win"; "The Dutch government believes herd immunity is more effective in fighting COVID-19"; etc.

We want to argue that the ontological sources of these facts are relational eventualities. Regarding collective agency, functionalism and interpretationism deal with two aspects of the relations across a collective, namely the internal relations of a collective (collective structure, that is, the interaction between its members) and the external relations of a collective (collective function, that is, the interaction between the collective and the outside world).

We believe that a collective agent is essentially a collection of relational eventualities. Specifically, this collection of relational eventualities includes relata (members of a collective and the targets of collective behaviors) and relations (a collective's internal structures and external functions). These relata and relations constitute a recognizable identity of a collective agent, which can then play a role in our decision-making. The irreducible collective intentionality is assigned to its members through internal relations, and an individual can choose to take part in the intentionality of different collectives depending on what net of relations they finds themselves in. For example, an agent could be in multiple

collectives simultaneously, such as their company, family, fishing club, sports team, etc. Each collective assigns collective intentionality to its members through its unique structure, and the individual-in-relation has to consider these collective intentions because they are involved in these collective relations.

This is a purely relational explanation. On the one hand, we do not presuppose any existence beyond what is commonly acceptable for our analysis, i.e., only substances and relations,¹⁶ without any mysterious element, such as a spiritual body. On the other hand, we reject physicalistic individualism and regard the existence of collective agency as a recognizable collection of relational eventualities. This view is consistent with the argument of Schmid [35] and Schweikard and Schmid [10], “It is, rather, to treat collective intentionality as irreducible with respect to its content and mode, and relational with respect to the structures that ground individuals’ reference to plural contents and their self-conception as subject-in-relations” [10].¹⁷

4.3. A Thorough Relational Account

Admittedly, some kinds of collective agents are not stable compared to individual agents, for example, a reading group. However, some kinds of collective agents look more stable than individual agents, for example, a university or a government. Either kind maintains its diachronic identity, the difference lies in the source of this maintenance. A collective with institutional structures or functions will not change by variation in members but only by variations in its structures or functions. We will not consider a basketball team a new team just because of its members’ changes. Nor will it be considered a new company just because the shareholders of a company are renewed. However, we will be intuitively aware of changes in collective identity due to changes in the internal and external relations of the collective. For instance, the internal structure has changed, such as the merger of a company with another one into a new enterprise, or the realization of the external functions can no longer meet the necessary conditions, for example, when it is revealed that the actual purpose of a so-called charity organization was to make a profit.

It seems that the relational patterns instead of the relata play a fundamental role in collective agency. The examples above show that a relata’s changing does not necessarily lead to a variation of our recognition of a collective, but a relation’s changing does the opposite. Can we speak of the relational pattern itself without relata? Of course! For instance, in modal logic, we can either use a concrete symbol Rab to express the relationship between states a and b , or we can use formulas such as T: $\Box\phi \rightarrow \phi$, 4: $\Box\phi \rightarrow \Box\Box\phi$, 5: $\Diamond\phi \rightarrow \Box\Diamond\phi$ to express more abstract relational properties. Similarly, our cognition of collective agency is actually more direct to its abstract relational patterns, and those relata only enter into relational eventualities by being referred to such relations.¹⁸

The relations in the collection of relational eventualities thus play a critical role in forming the fixed pattern of collective identity. Only when the pattern undergoes a fundamental transformation will we think that the collective has changed. It should be noticed that this does not mean we advocate a theory in which individual and collective agents are defined differently: a collective agent is determined by relations, whereas an individual agent is not. On the contrary, we hold that the concept of agency as such, not just that of collective agency, is essentially a collection of related eventualities that guarantee the realization of a diachronic rational perspective that can be applied and understood by others, and its actions satisfy necessary functional conditions. For an individual, this rational perspective is guaranteed by the stable and rational diachronic self-consciousness of the individual, while for a collective, this rational perspective is guaranteed by the relatively stable diachronic structure of the collective. Concerning functional conditions, we do not agree that the consistency of the system’s multiple decisions, as functionalism insists, is a necessary condition for the realization of rationality, but rather, more generally, that functional conditions merely ensure that actions embody the intentionality of the agent. In other words, even if multiple decisions of an agent are not consistent with each other,

as long as they are the result of the agent's intentionality, we still consider them to satisfy the functional conditions of agency.

We can re-examine the relationship between intention and agency from a relational perspective. In philosophy of action, intentionality is crucial because it is necessary for an agent to carry out their actions. From a relational point of view, intentionality represents the necessary relations that constitute agency; intentions and intentional actions directly bear the accessibility between cognitive contents and the external world. *Prima facie*, agency seems to be a property of an individual, but in essence, it is a relational property determined by the individual's various relational eventualities. The reason to say that agency is a relational property is that, in the broadest sense, agency expresses the ability to act as well as the capacity to manifest rationality, that presupposes and refers to the interactive relations between the agent and the external world (action), and indicates the nature of these relations to be manifested (rationality).

Thus, from a relational perspective, both individuals and collectives can be in relatively stable relational eventualities and thus can have the possibility to display agency. In terms of external functions, they have various interactions with the external world, and these interactions meet the same conditions, such as representation, motivation, rationality, etc. In terms of internal structures, individuals are fundamentally different from collectives. There are no inter-individual relations within individuals. Per contrast, a collective must have inter-individual relations as a collective structure to ensure its stability. On a relational view, these inter-individual relations are, from the perspective of an individual agent, between such an agent and the outside world since, for each individual, others categorically belong to the outside world too. However, from the perspective of collective agents, inter-individual relations constitute the stable internal relations of a collective. One may think of the internal relations of a collective as a subset of the union of various relations its members as individual agents are involved in. However, it should be noted that the properties of the same inter-individual relation from the individual perspective and from the collective perspective are different. The former is an accessible relation between cognitive content and the outside world, and the latter is part of the internal structure of a collective. Once collective intentionality has been formed, we can only regard the latter kind of inter-individual relations as part of internal structures, and not of relations of accessibility. This is the ontological reason at the micro-level why we have an irreducible concept of a collective in the context of collective intentionality. Moreover, besides those inter-individual relations that serve as the basic constitutive level of a collective's identity, a collective also contains higher-order relations, which take these inter-individual relations as *relata*. From a holistic perspective, together with these higher-order relations, all internal relations of a collective form a relational system that manifests the collective's identity. Thus, on such a holistic account, a collective identity grounded on its relational system cannot be adequately explained by its constituent parts. This is the ontological reason at the macro-level why we cannot reduce the concept of a collective. The question of how this collective structure comes into existence is beyond the scope of this paper. We assume that collectives by themselves have structures, and the discussion in this paper is within this boundary. Getting rid of this assumption will lead to more complex problems, a further relational account of agency left to future work.

Adopting a holistic perspective has some advantages: First, this is an entirely non-individualist account. We can fundamentally avoid the circularity problems faced by the orthodox theories and use the primitive concept of relation in explaining social phenomena. Second, holism can integrate interpretationism and functionalism into a unified theory and properly explain the irreducibility of the concept of a collective. Third, although this holistic theory recognizes a realism of collective agency, we have not introduced any Hegelian mysterious spiritual entities or any totalitarian specters that Popper has criticized. On the contrary, we only point out the individualist's excessive ontological demands and reaffirm the importance of relations as *bona fide*, acceptable ontological entities. Moreover, since

we have stated that the collective discussed here is structural and does not involve any spontaneous behavior, our account has a solid and defensible basis.

In conclusion, we say a collective has agency if and only if it has a relatively stable relational pattern to realize its rationality. Such a relational pattern includes at least one of the following relations:

- (1) A stable interrelationship between its members that constitutes a stable structure;
- (2) A stable relationship between its members' choices and the collective action that constitutes a stable procedure.

Any collective with agency that satisfies this condition will be counted as a collective agent. The relational pattern is critical because groups with no relatively stable relational pattern cannot manifest collective agency.

In the sense that the concept of a collective is presupposed in the conditions of an account, the relational account also involves analytical circularity. Nevertheless, the problem here is not severe because those relations in our conditions exist factually rather than conceptually, which means that the individual-to-collective transition that hampers individualist accounts presents no problem for ours. In the relational account, relations as sufficient and necessary conditions must exist in the ontological sense, which provides the source of the concept of a collective. Furthermore, by acknowledging the essential role of the relational pattern and its content, the relational account provides new knowledge about our understanding of collective agency, thus avoiding inferential circularity. The relational account combines interpretationism and functionalism and thus inherits their tradition of removing superfluous constraints on the concept of an agent and agency. In other words, in the relational account, as long as a group satisfies at least one type of the relatively stable relational pattern, we regard it as a collective agent possessing collective agency. This means that physicalism does not apply in the relational account. Furthermore, as mentioned above, acknowledging ontological relations leads the relational account away from individualism. Therefore, the problem of physicalistic individualism discussed in Section 3 is circumvented in the relational account.

5. Relations between Collective Agency and Individual Actions

There has always been a powerful objection to acknowledging the existence of collective agency: it is that collective agency will cause its members to lose their individual agency and become nothing but mechanical or instrumental units, which will constitute a major problem for rational theories that focus on individuals. However, this worry appears unfounded in a holistic account.

This objection reverses the order of one's feasible strategies and one's agency. In a holistic account, the agency of any system is manifested in its ability to evaluate, select and implement all its feasible strategies, and at the same time, reflect on its own rationality. This is not to say that an individual maintains their agency first, then participates in some collective agency and is restricted, i.e., collective agency rests on individual agency. This is the perspective that most analyses work from. Instead, in a relational network, individuals-in-relations show their agency by choosing between the optimal individual strategies and the optimal collective strategies.¹⁹ That is to say, not only does collective agency rest on individual agency, but individual agency also depends on its relation to relevant collective agency. There is a back and forth relationship between them, i.e., individual and collective agency are mutually dependent.

This does not mean that the choice of an individual agent can only be 0 or 1, i.e., to either manifest collective agency or their own agency. Note that in our definition of collective agency, a collective is a collection of relational eventualities. This introduces a temporal dimension. In a diachronic process, an individual agent can order the various relations they are involved in and gradually accomplish feasible strategies from different relations. For instance, someone finishes their work as a company's employee, then accompanies their child as a parent, then has dinner with their friend as part of maintaining a friendship, and then may go out cycling for their own pleasure. A variety of collective

agencies emerges in the diachronic actions of an individual agent, and individual agency is not compromised. Nevertheless, at any given time, if an individual chooses to implement the collective's optimal strategy, we believe that the collective achieves its agency. If an individual refuses to act in accordance with its status and functions in collective relations, we say the collective does not achieve its agency.

Another reason why the aforementioned objection is not valid is that participation in specific collective relations is always the result of an individual's intentionality. In daily life, individuals may function in multiple collectives. Different social relations bring optimal strategies in different dimensions; however, all things the agent has done are based on their own intentionality. They could choose to engage or not. Individual agency is not static. An agent is constantly updating their preferences and re-evaluating the relations they are in, and in that way they manifest diachronic agency. This is similar to the view of Rovane [37], "the existence of any agent is always a product of effort and will, and is never a metaphysical given—which is just to say, there are no natural persons." Our behaviors and choices regulate what kind of rational agent, what kind of person we really are.

The relational account also explains the phenomenon of sacrifice. If individual members only play an instrumental role in collective agency, then no one will be motivated to sacrifice individual gains. We would not say that those heroic figures who sacrifice for the collective benefit have no individual agency. On the contrary, this sacrifice manifests a kind of thinking that, after facing a variety of behavioral strategies and knowing all the consequences, the agent chooses to dedicate themselves to others in the collective. It is precisely because the individual agent evaluates all relevant relational eventualities and chooses to unify their own agency with their role in collective agency that other members feel respect for self-sacrifice and other forms of heroic behaviors.

On a relational account, the influence that collective agency will have on the behavior of its members is mainly divided into two aspects. On the one hand, members' behaviors will be affected in accordance with the particular structure a collective may have. If the collective has a flat structure, that is, all members have equal status in the collective, then the realization of collective agency requires each member to evaluate and decide among the feasible strategies involved in collective relations and ensure that the optimal strategy of the collective is chosen and implemented. If the collective contains a particular hierarchical structure, the hierarchy is reflected in its members' evaluations and decision-making. The subgroups that occupy the most central position should select strategies first regardless of other members, and then the next-level subgroups can choose based on the decision-making of the most central subgroup, and so on. In this sequential mode, the collective's optimal strategy selection is realized at different levels. Although, of course, collective structures in real society are more complicated in detail, it appears that these are the two most straightforward and common structures: either each level achieves collective optimality equally (small work team), or the choices of subgroups in different levels depending on the higher level together achieve the optimal collective strategy (large company). Changes in collective structures will correspondingly affect the decision-making sequence of its members. On the other hand, collective functions also affect the targets of collective actions realized by its members-in-relation. If relations between collective members and behavioral targets change, it will change the criteria of agency for collective actions. For example, changes in the industry a company is engaged in, such as shifting from furniture to grocery supply, will directly affect the regular operation of the whole enterprise in the short term and the specific work of its employees. In summary, changes in structures or functions of a collective substantially affect its members' actions.

6. Conclusions and Ideas for a Future Investigation

In this paper, we have reviewed several mainstream accounts of collective agency and collective intentionality and we have pointed out that they all manifest the tendency of physicalistic individualism resulting from long-term Cartesian influence. We have shown that the foundation of such physicalistic individualism is fragile and that theories

that subscribe to it cannot adequately explain the source of irreducible concepts of a collective. We then have developed a relational account, which regards the subject of collective intentionality as the relational identity composed of all the relative relational eventualities. An individual identifies and acts in a relational identity of a collective through their social relations, thus manifesting collective agency. Collective agency has both internal relations (collective structures) and external relations (collective functions). This confirms that interpretationism and functionalism about agency can be integrated into a relational, holistic account.

There is one kind of theorizing in the area of collective agency and collective intentionality that we have not said anything about in the paper so far, which is that of phenomenological interpretation. In addition, there is a trend that has gradually become notable, namely phenomenological interpretation. Starting from Husserl, many phenomenological interpretations of intentionality have been developed, and some of them have been integrated into contemporary accounts of collective intentionality. For example, Schmid [38] makes use of a phenomenological concept: pre-reflexive self-awareness can be applied to multi-agent situations. According to Schmid, an agent cannot engage in reflexive reflection without being aware of their own identity. Only based on pre-reflexive awareness can we consider our options for action and make decisions in our own situation without doubting our actions and ourselves. Therefore, when it applies to the collective level, collective agency will be considered as an adoption of pre-reflexive self-awareness of the collective identity. An investigation of the relation between this view and our relational account is, unfortunately, beyond the scope of this paper. Nevertheless, there is a possible alternative explanation of collective agency: the individual agent themselves are already able to recognize others as agents before their own pre-reflexive awareness (see Wittgenstein [39], a view that, arguably, finds support in cognitive psychological and anthropological research, as again in Tomasello [30]). Comparing these two viewpoints and further developing the relational account seems a promising direction for further research.

As mentioned in the introduction, all the discussions in this paper are based on two assumptions: a collective has structural characteristics, and we do not discuss spontaneous behaviors. Our relational account is successful with regard to structural collectives.²⁰ However, what if we broaden our horizons and discuss more situations, such as non-structural groups? For example, people who suddenly run together due to an alarm; who happen to pass by and offer assistance; and who walk together for no reason; etc. Various social phenomena need to be explained. For groups with opaque structures that do have constituted identities of collectives, it may be appropriate to talk about particular groups' pre-structural or quasi-structural characteristics on a looser basis.²¹

The relational, holistic account is still in its infancy, but its advantages are apparent enough. Physicalistic individualistic theories will be turned upside down from this perspective, forming a brand-new consistent interpretation of the three-body: individual–collective–society. As Karl Marx puts it, “but the human essence is no abstraction inherent in each single individual. In its reality it is the ensemble of the social relations” [40]. Combined with the relational account of collective agency, we will have a refreshed understanding of this sentence.

Author Contributions: Writing—original draft, Y.W.; Writing—review and editing, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Tsinghua University Initiative Scientific Research Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We want to thank Anthonie Meijers and Marc Slors for their comments on an earlier version of this paper, which have been helpful in further developing the argument. Of course, no agreement on their part is implied.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

- 1 Interpreting agency using the concept of intentionality is the mainstream tradition in philosophy of action since G. E. M. Anscombe and Donald Davidson.
- 2 The “rational” here indicates logical consistency of a collective’s decisions; that is, it is irrational for a collective agent to agree with A and B but to deny $A \wedge B$ at the same time. More detail on interpretationism and functionalism can be found in Section 4.
- 3 Following Bach [8], we use the term ‘eventuality’ to refer neutrally to events and states. Note that the reference to events or states instead of facts emphasizes the importance of the temporal dimension as a constitutive element. As a set of relational eventualities, collective agents can originate, develop and vanish. In other words, collective agents are not constants but relatively stable states of relations.
- 4 It is generally accepted that Wilfrid Sellars’ thought of we-intention (cf. [11]) was the source of Tuomela’s original analysis of we-intention (cf. [12]).
- 5 “This is a we-mode personal intention essentially of the kind ‘I, as a group member, intend to participate’” ([18], p. 44).
- 6 For instance, the example of business school ([22], pp. 47–48).
- 7 For the purposes of this paper, we will simplify the Cartesian claim by focusing only on internalism and leave out solipsism, i.e., we will take the existence of other individuals for granted.
- 8 Broadly speaking, in contrast to physicalism and reductionism, emergentism recognizes the existence of collective agency and asserts that such an entity can not be reduced to its members and their interactions. In this sense, interpretationism and functionalism can be regarded as specific ways of explaining emergentism by relaxing several restrictions on agency from their particular perspectives. However, in detail, their concerns are different from those of emergentism; they pay more attention to the way we interpret conceptual objects rather than whether they are reducible to other entities.
- 9 Here, we merely point out that there is an intuitive difference between knowledge of oneself and knowledge of the external world and that this difference affects the application of familiarity. It is beyond the scope of this paper to discuss in what sense this distinction holds. For interested readers, see [27] for various accounts of the distinction between the two types of knowledge.
- 10 In order for internalism to not collapse into solipsism, it needs to satisfy two conditions: 1. the subject regards perceptual inputs as evidence; 2. the subject has the criteria to distinguish between knowledge of oneself and knowledge of others. Therefore, the distinction between the two types of knowledge is still meaningful in the context of internalism, which affects the application of familiarity.
- 11 Many experimental studies in cognitive science and anthropology confirm that infants about 12 to 15 months of age are capable of having collaborative engagements. They can exercise their individual intentionality and recognize the intentionality of others, and more importantly, they can realize the cooperative behavior and the role of others in collaborative engagements and can take in and complete the tasks of other roles. (See [28,29]; etc.) With the observations that gorillas cannot have collective intentionality and that autistic children cannot fully have collective intentionality, Tomasello [30] further suggests that the most important distinction between humans and other species is that we understand collective intentionality and can cooperate, rather than merely being rational and intentional.
- 12 A common example is pain. Functionalists may interpret pain as a physical injury, a belief in physical discomfort, a desire to express such discomfort, moaning, restlessness, etc., and assume that anything that satisfies these conditions is capable of being pain. However, in this account, it is not just humans or animals that have internal mental states that are capable of pain, but also those that meet these conditions, even silicon-based creatures, aliens, and other completely different types, can be capable of pain. This is the so-called multiple realization: “pain can be realized by different types of physical states in different kinds of creatures” [31].
- 13 The “pattern” refers to the agential pattern in functionalism, which is the basic condition for behaviors of a system as an agent needs to satisfy. “There are purposes and representations that it is independently plausible to ascribe to the system... and the behavior of the system generally promotes those purposes according to those representations” [32].
- 14 For example, there are three propositions (a), (b) and (c) with interrelations: the conjunction of (a) and (b) implies (c), namely “(a) Take preventive measures against COVID-19”, “(b) Fighting the epidemic should have the highest priority, whatever it takes”, and “(c) Accepting financial loss.” According to functionalism, a collective that manifests agency can not simultaneously adopt (a) and (b) but reject (c) because it will be logically inconsistent and indicates the collective cannot exercise its ratiocinative capacity.
- 15 Donald Davidson and Daniel Dennett are prominent in this tradition.
- 16 Although some authors advocate ontological reductionism about relations, most philosophers acknowledge relations as ontological commitments. For detail, see [34].

- 17 One of the anonymous reviewers pointed out to us that a combination of organizational structures and functional capacity has already been made in the area of complex adaptive systems. In addition to convergence of opinions, in evolutionary economic theory emergence of generic rule structures plays an essential role. As a result, collective agents have internal processes that allow them to manifest rationality, independent of being observed as doing so. Interested readers can further check [36].
- 18 Admittedly, our understanding of modal operators and principles depends on a previous understanding of the meta-language definitions in terms of explicit relations and relata. The point here is that even when all the relata have been abstracted, we can still directly talk about the relational properties without referring to any concrete relata. Similar things happen when we talk about the essence of a collective.
- 19 In some cases, these strategies are inconsistent, such as the prisoner's dilemma.
- 20 Structures or functions of a collective guarantee minimum relational eventualities that constitute the relational identity of a collective.
- 21 The possibility of integrating our relationalist approach into the broader framework of critical realism, as suggested by one of the anonymous reviewers, is also a promising avenue to be explored.

References

1. Popper, K.R. *The Open Society and Its Enemies, 2 Vols., Reprint*; Princeton University Press: Princeton, NJ, USA, 1962.
2. Rovane, C. *The Bounds of Agency*; Princeton University Press: Princeton, NJ, USA, 1997.
3. Tollefsen, D. Organizations as true believers. *J. Soc. Philos.* **2002**, *33*, 395–410. [[CrossRef](#)]
4. List, C.; Pettit, P. *Group Agency: The Possibility, Design, and Status of Corporate Agents*; Oxford University Press: Oxford, UK, 2011.
5. Baier, A.C. Doing things with others: The mental commons. In *Commonality and Particularity in Ethics*; Springer: Berlin/Heidelberg, Germany, 1997; pp. 15–44.
6. Meijers, A.M. Can collective intentionality be individualized? *Am. J. Econ. Sociol.* **2003**, *62*, 167–183. [[CrossRef](#)]
7. Schmid, H.B. Can brains in vats think as a team? *Philos. Explor.* **2003**, *6*, 201–217. [[CrossRef](#)]
8. Bach, E. The Algebra of Events. *Linguist. Philos.* **1986**, *9*, 5–16.
9. Petersson, B. Collectivity and Circularity. *J. Philos.* **2007**, *104*, 138–156. [[CrossRef](#)]
10. Schweikard, D.P.; Schmid, H.B. Collective Intentionality. In *The Stanford Encyclopedia of Philosophy*, Fall 2021 ed.; Zalta, E.N., Ed.; Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2021.
11. Sellars, W. *Imperatives, Intentions, and the Logic of "Ought"*; Wayne State University Press: Detroit, MI, USA, 1963.
12. Tuomela, R.; Miller, K. We-intentions. *Philos. Stud.* **1988**, *53*, 367–89. [[CrossRef](#)]
13. Gilbert, M. *A Theory of Political Obligation: Membership, Commitment, and the Bonds of Society*; Oxford University Press on Demand: Detroit, MI, USA, 2006.
14. Gilbert, M. *Sociality and Responsibility: New Essays in Plural Subject Theory*; Rowman & Littlefield: Lanham, MD, USA, 2000.
15. Gilbert, M. Acting Together. In *Social Facts and Collective Intentionality. Philosophische Forschung/Philosophical Research*; Meggle, G., Ed.; Dr. Hänsel-Hohenhausen: Hesse, Germany, 2002.
16. Gilbert, M. *On Social Facts*; Princeton University Press: Princeton, NJ, USA, 1992.
17. Tuomela, R. *The Philosophy of Sociality: The Shared Point of View*; Oxford University Press: Oxford, UK, 2007.
18. Tuomela, R. Joint intention, we-mode and I-mode. *Midwest Stud. Philos.* **2006**, *30*, 35–58. [[CrossRef](#)]
19. Tuomela, R. We-intentions revisited. *Philos. Stud.* **2005**, *125*, 327–369. [[CrossRef](#)]
20. Tuomela, R. *Social Ontology: Collective Intentionality and Group Agents*; Oxford University Press: Oxford, UK, 2013.
21. Searle, J. Collective intentions and actions. *Intentions Commun.* **1990**, *195*, 220.
22. Searle, J. *Making the Social World: The Structure of Human Civilization*; Oxford University Press: Oxford, UK, 2010.
23. Bratman, M.E. Dynamics of sociality. *Midwest Stud. Philos.* **2006**, *30*, 1–15. [[CrossRef](#)]
24. Bratman, M.E. *Faces of Intention: Selected Essays on Intention and Agency*; Cambridge University Press: Cambridge, UK, 1999.
25. Bratman, M.E. *Shared Agency: A Planning Theory of Acting Together*; Oup Usa: Oxford, UK, 2014.
26. Humberstone, I.L. Two types of circularity. *Philos. Phenomenol. Res. Q. J.* **1997**, *57*, 249–280. [[CrossRef](#)]
27. Gertler, B. Self-Knowledge. In *The Stanford Encyclopedia of Philosophy*, Winter 2021 ed.; Zalta, E.N., Ed.; Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2021.
28. Bakeman, R.; Adamson, L.B. Coordinating attention to people and objects in mother-infant and peer-infant interaction. In *Child Development*; Wiley: New York, NY, USA, 1984; pp. 1278–1289.
29. Ross, H.S.; Lollis, S.P. Communication within infant social games. *Dev. Psychol.* **1987**, *23*, 241. [[CrossRef](#)]
30. Tomasello, M.; Carpenter, M.; Call, J.; Behne, T.; Moll, H. Understanding and sharing intentions: The origins of cultural cognition. *Behav. Brain Sci.* **2005**, *28*, 675–691. [[CrossRef](#)] [[PubMed](#)]
31. Levin, J. Functionalism. In *The Stanford Encyclopedia of Philosophy*, Winter 2021 ed.; Zalta, E.N., Ed.; Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2021.
32. Pettit, P. The reality of group agents. In *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*; Cambridge University Press: Cambridge, UK, 2009; pp. 67–91.
33. Pettit, P. Rationality, Reasoning and Group Agency. *Dialectica* **2007**, *61*, 495–519. [[CrossRef](#)]
34. MacBride, F. Relations. In *The Stanford Encyclopedia of Philosophy*, Winter 2020 ed.; Zalta, E.N., Ed.; Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2020.

35. Schmid, H.B. *Plural Action: Essays in Philosophy and Social Science*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009.
36. Page, S.; Miller, J.H. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*; Princeton University Press: Princeton, NJ, USA, 2009.
37. Rovane, C. Group agency and individualism. *Erkenntnis* **2014**, *79*, 1663–1684. [[CrossRef](#)]
38. Schmid, H.B. Plural self-awareness. *Phenomenol. Cogn. Sci.* **2014**, *13*, 7–24. [[CrossRef](#)]
39. Wittgenstein, L. *Philosophical Investigations*; John Wiley & Sons: Hoboken, NJ, USA, 2010.
40. Marx, K. Theses on Feuerbach (1845). In *Karl Marx and Frederick Engels, Selected Works*; Wildside Press: Rockville, MD, USA, 1970.