

On Help and Interpersonal Control

Emanuele Bottazzi and Nicolas Troquard*

LOA-ISTC-CNR, Trento, Italy

Abstract

Help is not much considered in the literature of analytic social philosophy. According to Tuomela (2000), when a helps an agent b (1) a contributes to the achievement of b 's goal, and (2) b accepts a 's contribution to the goal. We take a rather different tack. Our notion of help is one-sided and triggered by an attempt. It is one-sided because we can provide our help to someone without her accepting it. She could be unaware of our actions, or she could be unwilling to receive it. Helping is based on trying because it is agent b (supposedly) trying to do something that triggers a 's action of help. This is something supported for instance by Warneken and Tomasello's experiments with toddlers (Warneken and Tomasello 2006, 2009).

Help is interesting in its own right, but also because it allows us to reconsider the philosophical underpinnings of the essential notion of control in social philosophy. Help is seen here as a kind of weak interpersonal control, where an agent a 's agency guides an agent b 's agency.

When possible, we evaluate our framework on chosen scenarios taken from the literature in philosophy and psychology. The analysis is driven by a formal, logical approach. In particular, we make use of the modal logics of agency. This assists us in taking sensible philosophical choices, avoiding blatant inconsistencies. Moreover, the resulting formalism has the potential to serve as a computational engine for implementing concrete societies of cooperating autonomous agents.

1 Introduction

Helping behavior manifests itself in virtually every society. In fact, if collective action is an essential constituent of society, it may well be that helping behavior is a prerequisite ingredient of collective action. Instances of help in Human societies are "working as a hospital volunteer", "mailing off a charity donation to help hurricane victims", "cardiopulmonary resuscitation and rescue breathing on someone who has had a heart attack", etc. But helping behavior is commonplace in everyday interactions. It is not just a phenomenon occurring in emergency situations, or when somebody is in real need. There are also more trivial and common ways of helping others. For example "helping someone entering the metro by leaving room for them to get in", "helping

*✉ troquard@loa.istc.cnr.it

a child getting dressed”, “helping someone to gather some papers they accidentally dropped in a hallway”, etc. In the Encyclopedia of Social Psychology, it is defined as follows:

Helping behavior is providing aid or benefit to another person. It does not matter what the motivation of the helper is, only that the recipient is assisted. This is distinguished from the more general term prosocial behavior, which can include any cooperative or friendly behavior. It is also distinguished from the more specific term altruistic behavior, which requires that the motivation for assisting others be primarily for the well-being of the other person or even at a cost to oneself. ([24, p. 420])

The explanation of the reason for help is best left to social psychology. Although often focused on emergency situations, the study of decisions to help is a typical problem in the discipline. Latané and Darley proposed a decision model of helping [5]. Work of classifying helping behavior has also been done. Pearce and Amato [34] proposed a cognitively-based typology of helping along three dimensions: planned formal versus spontaneous informal; serious versus non serious; and giving or indirect versus doing or direct. Smithson and Amato [40] extended the classification with one dimension: personal versus anonymous.

If help has been a prominent topic of study in social psychology, the same cannot be said in philosophy. It is true that help is considered in ethics, but the typical questions that are explored there are: Is helping a duty? Are we required to help? Little, instead, has been written in analytic philosophy about what help is. We think that this is a loss, especially in the context of social philosophy. In the last years this stream of studies has been focused on the explanation of complex intertwinings of intentions and actions called *joint actions*. Typical scenarios under investigations are moving a sofa together [44], painting a house together [3], or preparing a hollandaise sauce together [39]. All these cases can be readily seen as the sum of some manifestations of help. Therefore we believe that an analysis of help itself may become important to tackle, in further studies, joint actions by means of it.

In our account an archetypical case of help—successful help—occurs when agent *b* tries to achieve a state of affairs, and *a* makes sure that, if *b* is trying to achieve some situation, then that very situation is the case. The contribution someone gives to the realization of that situation can vary. This means, for example, that we help others even if we don’t actively intervene into the situation: we see our partner trying to open the door and we help her by just seeing to it that she opens it. If she opens the door without us intervening, we helped her anyway since we, for example already reached the keys in our pocket, ready to open the door for her. As we shall see, this structure is a specialisation of a more general one. Help is a form of control over others’ agency. It is a way of monitoring what is going on and if necessary, provide what is needed to accomplish what the helpee is trying to accomplish. It is this preparedness to react as a backup-system that is the relevant part of helping behavior.

We will formalise a general concept of *weak* interpersonal control, a *guidance interpersonal control*, in the modal logics of agency commonly coined

“bringing-it-about”. See e.g. [21, 35, 22, 17, 6, 7, 13, 42]. It is a logic extending propositional logic with one modality E_i for every agent i . The formula $E_i\phi$ reads that “agent i brings about that ϕ ”, where ϕ describes some state of affairs. We will also make use of one modality A_i for every agent i , where $A_i\phi$ reads that “agent i tries to bring that ϕ ”. A first use of the attempt modality is probably due to Santos and others [37]. In the literature of the “bringing-it-about”, influence over agents has been subject to debate. One kind of *strong* interpersonal control—of agent a over agent b for ϕ —is simply captured by $E_aE_b\phi$.¹ More generally, it is any bringing about or attempt to bring about, by a of some conjunction where at least one conjunct concerns the agency of b : $X_a(X_b\phi \wedge \psi)$, where X_a and X_b is some modality of a ’s and b ’s agency respectively. In contrast, the pattern of weak interpersonal control will match $X_a(X_b\phi \vee \psi)$ (with ψ typically non-provably equivalent to the logical contradiction \perp). By instantiation of our general formalisation of weak interpersonal control, we will be able to discuss a variety of more specific controls, helps, and subjective helps. The logic will allow to express properties pertaining to helping behavior and reason about them rigorously. This will assist us in taking sensible philosophical choices. Moreover, the resulting formalism will have the potential to serve as a computational engine for implementing concrete societies of cooperating autonomous agents.

Control over a certain situation is central in Elgesem’s interpretation of the logics of “bringing-it-about” [6, 7]. Although the language is too abstract to discern all the nuances,² its proposed semantics at least offers a modelling guideline of agency in terms of Sommerhoff’s model of the goal-directed control that living things possess to achieve their function [41].

One of the main and yet somehow striking points of this kind of logic, is its “static” character. Actions are not considered along their temporal dimension. The notion of change, dynamics and time are abstracted away. Abstraction and modularity are the strengths of logic in general. It is because it abstracts away from some details of action makes “bringing-it-about” flexible and easily prone to modular upgrades [13].

It has recently been emphasized that “there is no *one* folk theory of action, in roughly the way there is no one folk tale of Little Red Riding Hood” [31, p. 91]. To us, the modal logic of “bringing-it-about” is very useful as a starting tool for the formal analysis of agency, and helping behavior in particular. Since philosophical and logical research on the notion of helping is in its pre-infancy, we believe that abstracting away from some details can be useful to discover at least some of its basic ingredients. As a logic of “doing”, “bringing-it-about” is indeed very apt to capture the essence of cases of *successful* interpersonal control. Successful cases are good starting points to explore tentative interpersonal control and helping behavior, as well as more “epistemic” cases. These are cases of being helpful but possibly ineffective. Trying replaces doing, and imperfect information brings in interesting troubles. Hence, the strength of the logic putatively lies in its very abstractness, as one can abstract away from distracting phenomena and still incorporate them

¹Some authors adopt the maxim *qui facit per alium facit per se* to emphasize that this strong interpersonal control implies full blown agency: $E_aE_b\phi \rightarrow E_a\phi$.

²The proposed axiomatisation in [7] indeed requires only simple minimal neighborhood models that are standard in modal logic. The axiomatisation was refined in [13], and proved complete with respect to a class of minimal models.

later ahead in the analysis.

2 Guidance interpersonal control

Logics of agency, and logics of “bringing-it-about” specifically, are the logics of the modalities E_x where x is an acting entity, and $E_x\phi$ reads “ x brings about ϕ ”, or “ x sees to it that ϕ ”. This tradition in logics of action comes from the observation that action is better explained by what it brings about. It is a particularly adequate view for *ex post facto* reasoning, and thus for discovering whether an acting entity is responsible at the moment of the achievement of an action. In a linguistic analysis of action sentences, Belnap and others [1, 2] adopt the *paraphrase thesis*: a sentence ϕ is agentive for some acting entity x if it can be rephrased as x sees to it that ϕ . Under this assumption, all actions can be captured with the abstract modality. It is regarded as an umbrella concept for direct or indirect actions, performed to achieve a goal, maintaining one, or refraining from one.

The philosophy that grounds the logic was carefully discussed by Elgesem in [6]. Suggested to him by Pörn, Elgesem borrows from theoretical neuroscientist Sommerhoff [41] the idea that agency is the actual bringing about of a goal towards which an activity is directed. Elgesem’s analysis leans also on Frankfurt [11, Chap. 6] according to whom, the pertinent aspect of agency is the manifestation of the agent’s guidance towards a goal. Sommerhoff’s goals are not necessarily goals proper, and instead are *telos* of an activity, that is, its terminus or end. This means that the notion of bringing about may refer also to non-intentional actions [16, 13] that have a final end anyway, related for example, to mere instinct.³

In Aristotelian terms, action (*praxis*) and production (*poiesis*) have, as their object, the contingent, that which can be otherwise (*to endechomenon allos echein*⁴). It is an important issue for whom is working in modal logic of agency. It was at the core of the discussions in early work such as [21, 22, 35], and in more recent examinations [17]. The crux of the issue is to capture the idea, in the semantics, that what the agent brings about has to be avoidable. In philosophy this is traditionally seen as *control*. To exercise control, possibilities have to be open to the agent. And this amounts to say that to bring about a state of affair ϕ is to exercise a control on ϕ . In [21, 22] this is linked to what is called *negative condition of agency*, that can be termed *counterfactual condition*, saying that if the agent had not acted the way she did, ϕ *would have not* been obtained [17]. The exact nature of this negative condition has been open to debate ever since. To mention only an eminent proposal, according to [35] this condition has to be weakened to the point that if the agent had not acted the way she did, ϕ *might have not* been the case.

With respect to that, Elgesem [7] makes an interesting point, holding that even this weaker negative condition is too strong. One can imagine cases where ϕ is the case, independently of what the agent does, but where it is still the case that he brings it about that ϕ :

³The main source is *Nicomachean Ethics* III. For a recent review on Aristotle’s voluntariness of action see [29].

⁴NE IV 5, 1140b 27

Consider this example. My one-year-old boy is in the process of learning to eat by himself. Sometimes he succeeds in getting the food into his mouth with the spoon, and sometimes not. Suppose he succeeds at some point during the meal, i.e. he brings it about that he has food in his mouth. During the whole of this meal, I am watching him to make sure that he gets fed. So if he does not succeed in getting the food into his mouth, I put the food into his mouth anyway. In this situation, it seems to be the case that there is no relevant alternative where it is not true that he gets food in his mouth. Now, in the case where he hits his mouth with the spoon, it must be right to say that he brings it about that he has food in his mouth. This is the case despite the existence of a reliable back-up system which guarantees that the goal is satisfied in any case.

In the context of the present study, the baby scenario is noteworthy for two reasons. The first one is that it suggests that we should consider a different notion of control. According to John Martin Fischer [8, 9], we can isolate at least two kinds of control: *regulative* control and *guidance* control. Regulative control is conceptually linked with the negative condition, because it requires freedom to choose and do otherwise. The notion of guidance control stems from Harry Frankfurt [10] and gives a better account of Elgesem's stance, because it does not require to consider necessarily that something can be otherwise. Guidance "is determined by characteristics of the actual sequence issuing in one's choice" [32]. Fischer, in order to illustrate this notion, proposes the example of driving a car. I have regulative control of the car if, given the fact that I wish to make a right turn, the car, as a result, moves to the right, but given the perfect condition of the car, I could have decided to make it turn to the left. Instead, suppose that the car is not in perfect condition, but has a quite peculiar malfunctioning such that, if I steer to the right it does it perfectly, but if I try and steer to the left, the car goes to the right, too. Suppose now that I actually steer the wheel to the right (the direction that does not display the malfunctioning). In this case I have guidance control of the car.

The second reason to find interest in Elgesem's scenario is that here, we are not simply dealing with agency, but with *interpersonal* agency. An interpersonal action, as justly observed by Seumas Miller [30], is an action that is *interdependent* with the action of some other agent, or is otherwise directed to an agent. In the case of the baby, the controlling agency is not just putting some food in a cavity, it is making sure that if the baby tries to put some food in his mouth, the food *is* in his mouth. The result may be realized with the contribution of his father, or by the baby actually feeding himself. We can call this *weak* or *guidance interpersonal control*, having two main components: the controlling agency and the controlled agency. The controlling agency is, in Elgesem's scenario, the father bringing about that *the baby is fed if the baby tries to be fed*.

In guidance interpersonal control, the controlling agency does not have "to go the way of" the controlled agency, though. Take for instance a case of counter-action. Imagine a rush-hour traffic scenario, and in particular these two cars side by side on two different lanes. When the driver in the car on the right lane (say Dr. R) tries to slot his car into the left lane, the driver in

the car on the left lane (say Dr. L) will accelerate ever so slightly to prevent it. The controlled agency is Dr. R's trying to have his car on the left lane. The controlling agency is here Dr. L bringing about that *Dr. R's car is not on the left lane if Dr. R tries to slot his car into the left lane*.

The notion of control that we want to highlight is a form of weak control indeed. Not only because interpersonal control is not regulative, but also because the control we are interested in is not a coercion of the controlled agent. It is not a form of constraining the agent into an unavoidable action, and it is not a form of mind control. It is control over a situation in which another agent is actively involved, and has an autonomously acquired volition. In Elgesem's son example the baby is not force-fed, but simply fed by his father when he tries and fails to do it by himself.

Trying, or attempt will become crucial in our work here. It has been analyzed in the philosophical literature, considered as a common feature of human actions and often linked to volition. (See [18, 33]. See also [26] for a review of the philosophical literature from a logical standpoint.) Trying clearly differs from effective agency. One can bring about something without even wanting it, for example by mistake. The *telos* of some bringing about is the final end of the action. It is in a way, where the action is directed, and it is not necessarily linked with volitions. When someone brings about that ϕ , we can say that ϕ is true. On the other end, when we consider the notion of trying, volition enters into the picture and from the fact that someone tries ϕ we cannot infer that ϕ is true. As highlighted also in the recent literature [19], one of the non obvious points related to trying is assessing whether someone has tried to do something whenever she has *succeeded* in doing it. We finally take no stance on the issue: we do not think that a bringing about implies an attempt, and we do not think that an attempt implies a bringing about. These principles will shine by their absence in the logic presented in the next section.

3 Logical aspects of guidance interpersonal control

3.1 The logic of "bringing-it-about" as a starting point

In this paper, we will use the logics of bringing-it-about (BIAT). It has been studied over several decades in philosophy of action, law, and in multi-agent systems ([23], [35], [25], [6], [37], [38], [7], [36], [12], [42]). It is the logic of the modality E_i , where i is an agent, and $E_i\phi$ reads " i brings about that ϕ ". Following [38], we will also integrate one modality A_i for every agent i , and $A_i\phi$ reads " i tries to bring about ϕ ".

We have laid out the main conceptual foundations of these operators in Section 2. We are now going to concentrate on the formal features of their logic.

Throughout the paper, we will assume a finite set of agents Agt and an enumerable set of atomic propositions Atm . The language of BIAT extends the language of propositional logic over Atm , with one operator E_i and one operator A_i for every agent $i \in \text{Agt}$.

The language L is defined by the following grammar:

$$\phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid E_i\phi \mid A_i\phi$$

where $p \in \text{Atm}$, and $i \in \text{Agt}$.

The fundamental principles (axioms schemes and rules and inference) of BIAT (where i is an individual agent) are:⁵

- (prop) $\vdash \phi$, when ϕ is a tautology of classical propositional logic
- (notaut) $\vdash \neg E_i \top$
- (success) $\vdash E_i \phi \rightarrow \phi$
- (aggreg) $\vdash E_i \phi \wedge E_i \psi \rightarrow E_i(\phi \wedge \psi)$
- (ree) if $\vdash \phi \leftrightarrow \psi$ then $\vdash E_i \phi \leftrightarrow E_i \psi$
- (rea) if $\vdash \phi \leftrightarrow \psi$ then $\vdash A_i \phi \leftrightarrow A_i \psi$

BIAT extends propositional classical logic (prop). An acting entity never exercises control towards a tautology (notaut). Agency is an achievement, that is, the culmination of a successful action (success). Agency aggregates (aggreg). Agency and attempts are closed under provably equivalent formulas (ree) and (rea).

We keep the logic of A_i very minimal. In particular, we do not take for granted that every actual agency requires an attempt. That is, we do not integrate $E_i \phi \rightarrow A_i \phi$ in the previous Hilbert system.

It is important to note that neither $E_i \phi \rightarrow E_i(\phi \vee \psi)$ nor $A_i \phi \rightarrow A_i(\phi \vee \psi)$ are derivable. They would indicate that agency and attempt are monotone modalities. We do not want that a bringing about that the letter is posted necessarily implies a bringing about that the letter is posted or the letter is burnt. In fact, adding the former formula to the axiomatization would yield an inconsistent theory. (In classical logic, it is incompatible with (notaut).) The logic of bringing it about is a weaker version of the achievement stit and of the deliberative stit in [2]. It is different from the Chellas' stit [20] which does admit the monotony of agency.

Strong interpersonal control. One typical kind of *strong* interpersonal control occurs when agent a brings about that an agent b brings about that ϕ . It is captured by $E_a E_b \phi$

More generally, a strong interpersonal control is any bringing about or attempt to bring about, by a of some conjunction where at least one conjunct concerns the agency of b . That is, where X_a and X_b is some modality of concerning a 's and b 's agency respectively:

$$X_a(X_b \phi \wedge \psi)$$

Decidability. The decidability of BIAT is important for its practical application in reasoning about social situations and procedures. The proof is a simple adaptation of the result obtained in [42].

Proposition 1. *Let a formula ϕ in the language of BIAT. The problem of deciding whether $\vdash \phi$ is decidable.*

This means that we can algorithmically decide of the validity of any property expressed in the language of BIAT. There is a procedure that one can mechanically follow, that will eventually provide the right answer to the question

⁵For any formula ϕ , the notation $\vdash \phi$ means that ϕ is provable within the logic. It is a theorem of the logic. That is, it is an axiom or a formula that can be deduced from the axioms and rules of inferences.

“is the formula ϕ valid?”, for every formula ϕ . There is a practical limitation in that the time complexity may grow exponentially with the size of the formula one wishes to automatically analyze. However, the task can be performed without an exponential blowup in space complexity.

The base logic is decidable but we do not claim so for the logics obtained by extending the above Hilbert system as suggested in the remaining of this paper. The problem for each single extension would require to be addressed individually.

3.2 The general form of guidance interpersonal control

We have seen in the previous sections that interpersonal control involves two interweaving actions: the controlled agency performed by a controlled agent b , and the controlling agency, performed by a controlling agent a . The latter capitalizes on the former to achieve some state of affairs, say γ . In order to logically characterize interpersonal control, we introduce three instrumental modalities, intended to capture *agentive modes*. We list them below along with a rough description of their purpose:

1. X_1^{ab} : used to capture the *mode of the controlling agency*;
2. X_2^{ab} : used to capture the *mode of the content of the controlling agency*;
3. X_3^{ab} : used to capture the *mode of the controlled agency*.

To reflect that the agentive modes are indeed modalities, we only need to assume that they obey the rule of equivalents:

$$\begin{aligned} (\text{rex1}) \quad & \text{if } \vdash \phi \leftrightarrow \psi \text{ then } \vdash X_1^{ab}\phi \leftrightarrow X_1^{ab}\psi \\ (\text{rex2}) \quad & \text{if } \vdash \phi \leftrightarrow \psi \text{ then } \vdash X_2^{ab}\phi \leftrightarrow X_2^{ab}\psi \\ (\text{rex3}) \quad & \text{if } \vdash \phi \leftrightarrow \psi \text{ then } \vdash X_3^{ab}\phi \leftrightarrow X_3^{ab}\psi \end{aligned}$$

The modalities must be expressible in the language but can take many forms. Example of modalities X that can be defined are $X\phi = E_i\phi$, $X\phi = E_i\neg A_j\phi$, $X\phi = E_j\neg A_i\phi$, $X\phi = A_j\phi \wedge E_i\neg\phi$, etc. For each example, it can indeed be readily checked that if $\vdash \phi \leftrightarrow \psi$ then $\vdash X\phi \leftrightarrow X\psi$. Despite this generality, we will frame more specifically the modalities intended to be used below.

Remark 1. *Instead of giving the modalities X_1^{ab} , X_2^{ab} , and X_3^{ab} a definition proper, we will use an axiomatic definition. For instance, instead of defining $X\phi = E_i\phi$, we would adopt $\vdash X\phi \leftrightarrow E_i\phi$ as an axiom. In such a way, we can flexibly provide partial, underspecified definitions of modalities. A weaker version of the previous example could be given as $\vdash X\phi \rightarrow E_i\phi$.*

The definition of a *guidance interpersonal control of agent a over agent b for γ* is then as follows:

$$\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \gamma) \stackrel{\text{def}}{=} X_1^{ab}(X_2^{ab}\gamma \rightarrow \gamma) \wedge X_3^{ab}\gamma$$

It is a general account for a situation, or state of affairs, describing a 's controlling agency over b 's agency, to obtain γ .

It is now better to progressively deconstruct $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \gamma)$. It is the general form of guidance interpersonal control and is a conjunction of two distinct states of affairs pertaining to some kind of agency:

- $X_1^{ab}(X_2^{ab}\gamma \rightarrow \gamma)$ is the controlling agency of the guidance interpersonal control;
- $X_3^{ab}\gamma$ is the controlled agency of the guidance interpersonal control.

In the controlled agency:

- X_3^{ab} is the agentive mode of the controlled agency.

In the controlling agency:

- X_1^{ab} is the agentive mode of the controlling agency;
- $X_2^{ab}\gamma \rightarrow \gamma$ is the content of the controlling agency;
- X_2^{ab} is the agentive mode of the content of the controlling agency.

Typically then, we will think of the modalities reflecting more specific modes than suggested before. The modality X_1^{ab} would reflect some agentive mode pertaining to a . To commit the definition to a more definite flavor of *control* of a , we will consider that X_1^{ab} is either A_a or E_a . Practically, it means that we will only consider such instantiations in this paper. Agent a 's control is over b 's agency. So in the instantiations of guidance interpersonal control considered in this paper, the modalities X_2^{ab} and X_3^{ab} will always reflect some agentive mode pertaining to b . The main idea is that (i) the controlled agency indeed reflects b 's agency, (ii) the controlling agency indeed reflects a 's agency, and (iii) the content of the controlling agency partly reflects b 's agency.

The rush-hour traffic scenario. Remember Dr. R trying to slot his car into the left lane, and Dr. L making sure that it does not happen if he does try. Take left to mean that Dr. R's car is on the left lane. The guidance interpersonal control at play in the scenario can be instantiated as follows:

$$\text{GIC}(E_{DrL}, A_{DrR} \neg, A_{DrR} \neg, \neg\text{left})$$

which translates into:

$$E_{DrL}(A_{DrR}\text{left} \rightarrow \neg\text{left}) \wedge A_{DrR}\text{left}$$

3.3 Some formal properties of interpersonal control

As a general definition, our formal account of guidance interpersonal control (of a over b) can be instantiated to specific cases by simply identifying the three agentive modes to a particular modality expressible in the logic of BIAT. We can then use the formal tools provided by the logic to rigorously define a terminology pertaining to the properties of interpersonal control. We begin with a few simple qualities.

An interpersonal control is *well-situated* when the agentive mode of the controlled agency coincides with the agentive mode of the content of the controlling agency.

Definition 1 (WS). For any ψ , a guidance interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \psi)$ is well-situated when

$$\vdash X_3^{ab}\psi \leftrightarrow X_2^{ab}\psi$$

Intuitively, well-situatedness is a good property for a 's controlling agency. Indeed, in a well-situated controlled agency (for γ), γ is true iff the content of the controlled agency is true.⁶ With the right mode of controlling agency, a 's agency can then be the least effort for γ .

What more sure way to have the content of a well-situated controlled agency true, and hence γ , than to effectively bring it about? An interpersonal control is *effective* when the agentive mode of the controlling agency coincides with a bringing about of agent a .

Definition 2 (EF). For any ψ , a guidance interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \psi)$ is effective when

$$\vdash X_1^{ab}\psi \leftrightarrow E_a\psi$$

Uncertain, unskilled, or hazardous controlling agency by a would remain a worthwhile effort. We say that an interpersonal control is *tentative* when the agentive mode of the controlling agency coincides with an attempt of agent a .

Definition 3 (TE). For any ψ , a guidance interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \psi)$ is tentative when

$$\vdash X_1^{ab}\psi \leftrightarrow A_a\psi$$

Remember that we framed earlier X_1^{ab} to be identified either as A_a or as E_a . No other mode of controlling agency will be considered here. In this context, definitions 2 and 3 offer a clear dichotomy of guidance interpersonal control: (i) effective control, which is actual and successful,⁷ (ii) tentative control, which is an uncertain, possible control.

These properties of interpersonal control are presented as provable logical formulas in the language of BIAT. Methodologically, it means that to design a logic of guidance interpersonal control, it suffices to combine into a Hilbert system:

1. the axiomatic system of BIAT presented on Page 7;
2. the principles (rex1), (rex2), and (rex3);
3. the definition of $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \gamma)$;
4. a set of properties of guidance interpersonal control.

In Section 5, we will address a few properties of interpersonal control that are more specific to the notion of help, and we propose a few simple theorems to exemplify the kind of reasoning that is enabled by our formal proposal.

4 Help

Help is a form of *weak, interpersonal, guidance control of agent a over agent b for some state of affairs ϕ* . It goes, so to say, in the way of the helpee's trying or attempting. The controlling agency here is for the sake of the controlled one. In order to provide help, some control over the situation in which the

⁶We have $\vdash (\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \gamma) \wedge (X_3^{ab}\gamma \leftrightarrow X_2^{ab}\gamma)) \rightarrow ((X_2^{ab}\gamma \rightarrow \gamma) \leftrightarrow \gamma)$.

⁷Successful in virtue of axiom (success).

helped is trying to achieve some ϕ is needed. The baby example provided in Section 2 is not just an example of interpersonal guidance control, but it is also an example of *help*. The father is helping his child to get fed, and this does not mean that all the time he is materially putting the food in his mouth. The father is exercising guidance control of a disjunctive state of affairs involving the child trying to get fed. If the child is able to get fed by himself, there is no need to intervene. On the other hand, if the child tries but fails, then the father's agency will have him intervene in guiding the food in the mouth of the baby. In contrast, the rush-hour traffic example (Section 2, and end of Section 3.2), although a weak interpersonal guidance control, certainly is not a helping behavior.

As we said the notion of help is not much studied in philosophy in its own right. A notable exception, Raimo Tuomela [43] endorses that helping is in essence *a* adopting *b*'s goal and *b* accepting it. It is then a special case of asymmetric cooperative activity. If *b* has much to carry and *a* has no load, *a* may offer to help *b* to carry some of *b*'s bags. In thus helping *b*, *a* engages in a cooperative activity with *b* and *b* accepts *a*'s help:

a helps *b* relative to *b*'s autonomously acquired goal to achieve γ if and only if a) *a* intends to contribute to *b*'s achieving γ and carries out this intention, and b) *b* accepts a). [43, p. 136; adapted notation].

It is important to point out at once that Tuomela himself sees his characterization as too strong [43, p. 136]. We are seeking specifically a more basic, and, at the same time, more general notion. We would like to contrast it with the two conditions provided by Tuomela.

First, let us consider the point of view of the helper *a*, that is the a) condition of Tuomela's definition. As we have already said, the emphasis in our framework is not on the intentional notion of goal of an agent, it is rather on the more general notion of end (*telos*) of an action. With this in mind, not to be committed to strictly purposeful actions can also leave room for helping behavior in other forms of agency. For example, what appears to be a spontaneous tendency of children to cooperate [46] could be seen as an impulsive helping behavior:

The behavior is as simple as it is surprising—and it is highly robust. Drop an object accidentally on the floor and try to reach for it, for example, from a desk, and infants as young as 14-18 months of age will toddle over, pick it up and return it to you. [46, p. 397].

This obviously depends on what position one may take with respect to intentions. If impulses are considered as intentions, then Tuomela's definition is valid, in this respect at least. If, instead, we are not willing to accept impulses as some form of intentions, then our teleological notion of help is more flexible, since it covers both options. But there are other cases that do not fit Tuomela's definition in any way. Consider, for example, a competitive game, where some unintentional behavior of some player just helped the opponent in taking advantage in the game. For instance, a football kicked by an attacker and bouncing off a defender into the net. Out of all the possible positions on the field, the defender chose this one. It is a controlling agency of the defender

that is ill-fatedly directed towards making sure that the opponent's attempt to score is realized. Finally, one can imagine cases of help also in actions where it is difficult to assess if they are intentional or unintentional, as in side effects or lucky actions [28]. These cases exclude to us as requirements both the adoption of someone's goal and the formation of an intention. In such cases the teleological stance that we adopt shows instead its benefits.

Secondly, the other point of view to consider is the helpee's one. The first condition regarding b is that b 's goal has to be "autonomously acquired". This assumption is meant, as Tuomela himself states, to exclude cases where a coerces b to have the goal γ and b accepts the "help" in virtue of that coercion. This condition seems significant also in the light of what we said about weak interpersonal control, that also applies to our conception of help. As we stated in Section 2, the interpersonal control we are interested in is neither a form of mind control nor some way of bringing about that the helpee brings about that something is the case. This amounts to say that the control provided in a helping behavior has to be over a conditional state of affairs, whose antecedent is a proper volition/trying of the helpee. (Classically, this conditional is also a disjunctive state of affairs where one disjunct is the negation of the trying.) The relevance of the helpee's rational volition is also the primary assumption taken by Chisholm and Zimmerman (in an otherwise mysterious working note):

My being helped by someone to bring about some event implies an intentional relation between me and the event in question. Jones's helping Robinson to do something implies that Robinson, at least, "knows what he's doing", whether or not Jones does. [4, p. 402]

The other condition imposed by Tuomela that regards b , the acceptance condition—that is, the requirement that b accepts that a intends to contribute to b 's achieving γ and carries out this intention—is instead more problematic. There are many cases of help where the acceptance is not needed, because the controlling agency fits, so to say, with the volition of the helpee, with no need for the helpee to have any agreement on it. Consider cases of paternalistic help, that is a rather common manifestation of help in human behavior. Recent studies in developmental psychology show how when facing the situation where an experimenter requests something that is ill-suited to achieving their ultimate goal, 3-year-old children override the request in favor of what they believe is best for them [27]. There is no acceptance from the helpee and yet, help occurs. The same goes in our previous football example where the defender unwittingly helps the attacker to score. The attacker is not expected, in order to be successfully helped, to accept what the unlucky defender's agency is going to provide him. Even if the helpee is unaware of the helper's agency, it is sufficient for him to take advantage of the situation, and the resulting event can be considered help.

Given these observations we can now focus on our notion of *successful help*:

a helps b relative to b's trying to bring about that ϕ , if and only if: (i) a brings it about that: if b tries to bring about that ϕ then ϕ is the case; (ii) b tries to bring about that ϕ .

Condition (i) represents the controlling agency of help. It is a bringing about, so it is an *effective* guidance interpersonal control. Its content is a ma-

terial implication, dependent on the helpee's volition. As the left hand side of the condition is exactly b 's attempt to bring about ϕ , we will qualify it as a *justified assistance*. We can have a successful case of help when this controlling agency properly combines with condition (ii), the controlled agency. It requires that the helpee has to actually try to bring about that ϕ . We will qualify this property as an *opportune assistance*. As these two agencies are properly aligned, we are in presence of a *well-situated* interpersonal control, formally defined in the previous section. It implies that help is *successful* and ϕ is the case.

Since the controlling agency is conditional, agent a can help b without necessarily actively intervening in the situation. Agent a 's agency may be decisive for the truth of ϕ or may be redundant. We will define in particular the fact that the assistance is *decisive* when b does not bring about ϕ himself.

Elgesem's example is exactly about this. He assists his son for the sake of his son's attempt to be fed. If the baby is able to do it by itself the assistance is not decisive, otherwise, Elgesem makes sure that the baby gets fed when it tries. The example of the keys mentioned in the introduction, is also in line with our definition. If we see our partner about to open door, as we reach for the keys in our pocket ready to open the door, we help her anyway, even if she does get the keys first and does open the door. Help is exactly about the preparedness to provide to the helpee, what is needed in order to accomplish what she is trying to accomplish. It is also what Elgesem calls a back-up system in his example. The preparedness is a kind of guidance interpersonal control. But, we want to emphasize it, this control is *weak*. First, it is not mind control or some other strong way to bring about that the helpee brings about that ϕ holds. Agent b can autonomously acquire her volition $A_b\phi$. Second, it is a form of guidance. The state of affairs ϕ could become true even without an active participation of the helper. And even if his intervention is decisive, ϕ has to become true only with the precondition of the helpee's volition.

5 Logical aspects of help

5.1 Some formal properties of assisting behavior

On our way to characterize the notion of successful help that we defended before, we propose a few properties pertaining to what we may call more generally *assisting behaviors*, or simply *assistances*.

To start with, paternalism is a limiting factor to the meaningfulness of help. It occurs when the controlling agency of a does not properly capitalize on an attempt of b to bring about γ . We then start by characterizing a condition for the controlling agency to be a justified assistance (of a towards b).

Definition 4 (JA). For any ψ , a guidance interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \psi)$ is a justified assistance when

$$\vdash X_2^{ab}\psi \rightarrow A_b\psi$$

We say that an interpersonal control is a justified assistance when the mode of the content of the controlling agency at least includes an attempt of b .

All assistances are not necessary for bringing about the state of affairs sought after by a controlled agency. One class of these assistances is that of faked assistances. They occur when the mode of the content of the controlling agency at least includes b bringing about its volition.

Definition 5 (FA). For any ψ , a guidance interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \psi)$ is a faked assistance when

$$\vdash X_2^{ab}\psi \rightarrow E_b\psi$$

So the controlling agency of a faked assistance is over a state of affairs satisfying γ when, at least, b does bring about γ .

Critical properties of interpersonal control depend on the controlled agency. A decisive assistance occurs when the controlled agency of an interpersonal control for γ implies that b does not already bring about γ .

Definition 6 (DA). For any ψ , a guidance interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \psi)$ is a decisive assistance when

$$\vdash X_3^{ab}\psi \rightarrow \neg E_b\psi$$

This is decisive in the sense that b does not bring about γ himself. This is not necessarily decisive in the sense that γ would not be true if it were not for b 's action. Indeed, γ might be true coincidentally for some reason independent of a and b 's actions. One can of course define a stronger property as follows:

Definition 7 (SD). For any ψ , a guidance interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \psi)$ is a strongly decisive assistance when

$$\vdash X_3^{ab}\psi \rightarrow \neg\psi$$

Surely however, assistances would barely deserve the name if it were not for b to actually try to bring about a state of affairs.

Definition 8 (OA). For any ψ , a guidance interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \psi)$ is an opportune assistance when

$$\vdash X_3^{ab}\psi \rightarrow A_b\psi$$

In an opportune assistance for γ , the controlled agency at least implies that b tries to bring about γ .

5.2 A simple account of successful help

Finally, successful help (of a towards b for γ) can be rigorously defined as the weakest form of well-situated opportune effective assistance. That is, $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \gamma)$, where:

1. $\vdash X_1^{ab}\psi \leftrightarrow E_a\psi$
2. $\vdash X_2^{ab}\psi \leftrightarrow A_b\psi$
3. $\vdash X_3^{ab}\psi \leftrightarrow A_b\psi$

It is worth defining a new dedicated modality. Thus, we obtain:

$$[a : b]\gamma \stackrel{\text{def}}{=} E_a(A_b\gamma \rightarrow \gamma) \wedge A_b\gamma$$

which we read “*a* successfully helps *b* to bring about γ ”.

It is *successful* because we have the following expected property by applying (success) and (prop):

Proposition 2. $\vdash [a : b]\phi \rightarrow \phi$

It is an assistance for three reasons. First, there is an *assiste*. It is a volition of *b* to bring about γ and *b* does try. Second, there is a *assistant*. The content of *a*'s control is over the state of affairs where γ is true whenever *b* tries to bring about γ . Hence, *a*'s guidance is reactive to *b*'s goodwill in the action. Third, it is compelling to a formalization of assistance that $[a : b]\gamma \wedge \neg E_a\gamma \wedge \neg E_b\gamma$ is a consistent formula. That is, it is possible that *a* successfully helps *b* to bring about γ , and still, neither *a* nor *b* brings about γ . Hence, the success of the assistance described by $[a : b]\gamma$ comes from some cohesion between *a* and *b*.

Elgesem's example. Back to Elgesem's example about his one-year old boy (see Section 2). There are two cases: “sometimes [the boy] succeeds in getting the food into his mouth with the spoon, and sometimes not.” When he does succeed, Elgesem argues that the boy does bring about that he has food in his mouth. That is, $E_{\text{boy}}\text{food}$, where *food* stands for “the boy has food in his mouth”. When he does not succeed however, there is a “back-up system”. It is, we argued, the help provided by the father. Note that the accent is put on the boy being in the process of learning to eat by himself. There is no case of feeding the boy against his will. So, we must say that indeed the boy tries to bring about that he has food in his mouth: $A_{\text{boy}}\text{food}$. It is the controlled agency. The back-up system is the controlling agency, which consists in making sure that the boy has food in his mouth when he tries to bring about that the food is in his mouth. The agent of the controlling agency is Dag Elgesem himself, so we have: $E_{\text{dag}}(A_{\text{boy}}\text{food} \rightarrow \text{food})$. This is a case of effective, well-situated opportune guidance interpersonal control of Dag over the boy's agency towards the boy having food in his mouth.

To sum up, at least one of the following holds:

- $E_{\text{boy}}\text{food}$
- $E_{\text{dag}}(A_{\text{boy}}\text{food} \rightarrow \text{food}) \wedge A_{\text{boy}}\text{food}$

which implies that *food* holds no matter what.

5.3 Proven properties of interpersonal control and assistances

The logical theory allows to reason about more complex properties of contextual agency now expressible with our vocabulary. Some properties are expected from the choice of terminology. We can verify for instance that a strongly decisive assistance is a decisive assistance.

Theorem 1. *Strongly decisive assistance is decisive.*

PROOF.

1. $\{SD\} \vdash X_3^{ab}\gamma \rightarrow \neg\gamma$ (from SD)
2. $\{SD\} \vdash \neg\gamma \rightarrow \neg E_b\gamma$ (from (success) and (prop))
3. $\{SD\} \vdash X_3^{ab}\gamma \rightarrow \neg E_b\gamma$ (from 1., 2., and (prop))
4. $\{SD\} \vdash DA$ (from 3. and DA)

But typically, properties are not so transparent. We prove two more theorems.

Theorem 2. *Opportune well-situated assistance is justified.*

PROOF.

1. $\{OA, WS\} \vdash (X_3^{ab}\gamma \rightarrow A_b\gamma) \wedge (X_3^{ab}\gamma \leftrightarrow X_2^{ab}\gamma)$ (from OA, WS, and (prop))
2. $\{OA, WS\} \vdash X_2^{ab}\gamma \rightarrow A_b\gamma$ (from 1. and (prop))
3. $\{OA, WS\} \vdash JA$ (from 2. and JA)

Theorem 3. *Effective faked assistance is impossible.*

PROOF.

1. $\{EF, FA\} \vdash X_1^{ab}\gamma \leftrightarrow E_a\gamma$ (from EF)
2. $\{EF, FA\} \vdash X_2^{ab}\gamma \rightarrow E_b\gamma$ (from FA)
3. $\{EF, FA\} \vdash X_2^{ab}\gamma \rightarrow \gamma$ (from 2., (success), and (prop))
4. $\{EF, FA\} \vdash E_a(X_2^{ab}\gamma \rightarrow \gamma) \rightarrow \perp$ (from 3., (notaut), and (prop))
5. $\{EF, FA\} \vdash X_1^{ab}(X_2^{ab}\gamma \rightarrow \gamma) \rightarrow \perp$ (from 1., 4., and (prop))
6. $\{EF, FA\} \vdash X_1^{ab}(X_2^{ab}\gamma \rightarrow \gamma) \wedge X_3^{ab}\gamma \rightarrow \perp$ (from 5. and (prop))
7. $\{EF, FA\} \vdash \text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \gamma) \rightarrow \perp$ (from 6. by definition)

In English: effective faked assistance is impossible because it occurs when (i) the agentive mode of the assistant is to actually bring about the content of the interpersonal control (for γ), and (ii) the controlled agency includes the fact that the assistee already brings about γ . But by (ii) and (success), the content of the controlling agency is trivial: it is a theorem in the logic. But by axiom (notaut), the logic does not allow an agent to bring about tautologies, which is what the assistant's mode is by (i).

On the other hand, tentative faked assistance is possible. The reason is rather ordinary: according to our axiomatics of BIAT, it is possible for an agent to attempt to bring about tautologies.

6 Subjective help

We have been arguing for and formalizing an account of help which is unilateral and triggered by an attempt. It is unilateral because we can provide our help to someone without her accepting it. She could be unaware of our actions, or she could be unwilling to receive it. Help is based on trying because it is agent b (supposedly) trying to do something that triggers a 's action of help.

Here, we want to add that subjectivity plays a crucial role for characterizing an event as an event of help. Help is subjective since in helping b , agent a can have imperfect information about b 's volition. There was a Norwegian TV commercial for *Japp* chocolate bars where a man finishes a jog on a mountain road and arrives panting at his sport car parked near a cliff. He proceeds to stretch, hands on the car, facing the cliff. With a background of Caribbean music, a rastaman is driving by, eating a chocolate bar. (The slogan says that it gives extra energy.) He sees the scene, and looking determined he stops his truck, jumps out, walks to the car and pushes it over the cliff. As the rastaman believed that the car owner was trying to push his car off the cliff, there is an aspect of helping behavior in this event.

We extend the BIAT framework with one modality Bel_i for each agent i . The formula $Bel_i\phi$ reads that the agent i believes ϕ . Since our basis framework of agency is very abstract (BIAT is a weak modal logic), we do not assume much about the logic of Bel_i .

Any logic between S5 (full blown knowledge [14]), and the minimal modal logic should be consistent with our analysis in this paper. (Intermediate systems can be found in [15].) Although we will not do specific reasoning about beliefs in this paper, it is typically judicious for a work in modal logic to assume the following:

$$(reb) \quad \text{if } \vdash \phi \leftrightarrow \psi \text{ then } \vdash Bel_i\phi \leftrightarrow Bel_i\psi$$

We need a feasible methodology to pick out *events of subjective help* out of the many types of weak interpersonal control. We must concede that we cannot think of a unique methodology that would explain satisfyingly and completely why we consider that some instance of interpersonal control is not an event of help and why we consider some other instance as a typical event of help. Nonetheless, we can reiterate what aspects we see as relevant, propose the pertinent sets of parameters (viz., X_2^{ab} and X_3^{ab}) and exhaustively analyze their possible combinations.

The relevant aspects of subjective help are:

- it is based on a (presumed) attempt on the assistee part;
- it is subjective on the assistant part.

Assuming that the relevant beliefs of the assistant concern the trying events of the assistee, this considerably restricts our research space. Finally, we only consider help as an effective agency. Thus we adopt EF, meaning that X_1^{ab} has to be E_a . We will later comment on replacing effective agency by attempt agency.

Identifying the relevant subjective events of effective help. With the previous considerations in mind, we will look specifically at the cases of interpersonal control $\text{GIC}(X_1^{ab}, X_2^{ab}, X_3^{ab}, \gamma)$ where the mode of the controlled agency X_3^{ab} and the mode of the content of the controlling agency X_2^{ab} can obey one of three possible principles. We will examine the following (for all nine combinations of $X = X_2^{ab}$ and $X = X_3^{ab}$):

- $\vdash X\psi \leftrightarrow A_b\psi$
- $\vdash X\psi \leftrightarrow \text{Bel}_a A_b\psi$
- $\vdash X\psi \leftrightarrow A_b\psi \wedge \text{Bel}_a A_b\psi$

For clarity of exposition we will use several variations on a toy scenario of interaction between agent a and agent b , where a operates two push-buttons 1 and 2, and b operates a push-button 3. A light is on (property captured by γ) iff 1 is pressed, and at least one of 2 and 3 is pressed. Suppose that only agent b may have some concern over γ , and pushes his button 3 as a way to try to bring about that the light is on: $A_b\gamma$. Agent a can assist b in doing so, but may have imperfect knowledge as to whether b indeed tries to bring about γ . Either a believes that b tries to bring about γ ($\text{Bel}_a A_b\gamma$) or does not ($\neg \text{Bel}_a A_b\gamma$).

When the mode of the content of the controlling agency is A_b , the interpersonal control is (minimally) justified. In the lights of our toy scenario, the controlling agency may be seen as a indiscernibly pressing button 1, no matter what his beliefs are. If b tries to bring about γ , thus pressing the button 3, γ would hold.

- $E_a(A_b\gamma \rightarrow \gamma) \wedge A_b\gamma$. It is precisely our account of successful help: effective, opportune and well-situated interpersonal control.
- $E_a(A_b\gamma \rightarrow \gamma) \wedge \text{Bel}_a A_b\gamma$. It is not (necessarily) an opportune assistance. It also does not ensure that γ indeed holds. Agent a believes that b tries to bring about γ , but this belief is not taken into account in the controlling agency.
- $E_a(A_b\gamma \rightarrow \gamma) \wedge (A_b\gamma \wedge \text{Bel}_a A_b\gamma)$. It is logically equivalent to the conjunction of the two previous cases. It is an effective, opportune and well-situated interpersonal control, and agent a 's belief does not add anything remarkable.

When the mode of the content of the controlling agency is $\text{Bel}_a A_b$ we face a subjectively sensitive case of controlling agency. It is not justified. In our scenario, the controlling agency may be seen as a pressing the button 1 no matter what, and also pressing 2 whenever he believes that b tries to bring about γ .

- $E_a(\text{Bel}_a A_b\gamma \rightarrow \gamma) \wedge A_b\gamma$. Although it is an opportune assistance, it does not (necessarily) imply that γ holds.
- $E_a(\text{Bel}_a A_b\gamma \rightarrow \gamma) \wedge \text{Bel}_a A_b\gamma$. It is not (necessarily) an opportune assistance but γ holds.

- $E_a(Bel_a A_b \gamma \rightarrow \gamma) \wedge (A_b \gamma \wedge Bel_a A_b \gamma)$. It is logically equivalent to the conjunction of the two previous cases. It is an opportune assistance, and the agent a 's belief has the effect that the interpersonal control results in γ being true.

Remark 2. *We can observe that our description of the variants of the toy example suggests that in the previous second and third cases a presses both push-buttons 1 and 2. For all practical purpose we might say, in this example, that a does bring about that γ . We prefer to leave the question open in this paper whether it should be a general principle. . . Possibly, it could be argued that $(E_a(Bel_a A_b \gamma \rightarrow \gamma) \wedge Bel_a A_b \gamma) \rightarrow E_a \gamma$ would make a pertinent principle of agency.*

When the mode of the content of the controlling agency is $A_b \psi \wedge Bel_a A_b \psi$, the interpersonal control is justified, and the controlling agency is subjectively sensitive. In the scenario, the controlling agency may then be seen as the variant where a presses the push-button 1 whenever he believes that b tries to bring about γ .

- $E_a((A_b \gamma \wedge Bel_a A_b \gamma) \rightarrow \gamma) \wedge A_b \gamma$. It is an opportune assistance, but it is not (necessarily) true that γ .
- $E_a((A_b \gamma \wedge Bel_a A_b \gamma) \rightarrow \gamma) \wedge Bel_a A_b \gamma$. It is not (necessarily) an opportune assistance, and it is not (necessarily) true that γ .
- $E_a((A_b \gamma \wedge Bel_a A_b \gamma) \rightarrow \gamma) \wedge (A_b \gamma \wedge Bel_a A_b \gamma)$. It is logically equivalent to the conjunction of the two previous cases. It is an opportune assistance, a justifiably believes that b tries to bring about γ . It does imply that γ holds.

Tentative subjective help. Each case of effective help that we just mentioned naturally has a counterpart as tentative help.

In order to talk conveniently about tentative subjective help, we must come up with an adequate modification of the toy scenario used previously. Agent a now is at some distance from the push-buttons 1 and 2, and has to throw skillfully a juggling ball at each of them in order to activate them. What is important here is that unlike pushing a button, the result of throwing a ball at a button has a non-deterministic result. Throwing a juggling ball at a button, we consider it as a trying to press the button. Agent b still operates the push-button 3, normally as before. In addition, the light is on in the same conditions as before, that is, when 1 is pressed, and at least one of 2 and 3 is pressed.

Consider again the three cases:

- $\vdash X_2^{ab} \psi \leftrightarrow A_b \psi$
- $\vdash X_2^{ab} \psi \leftrightarrow Bel_a A_b \psi$
- $\vdash X_2^{ab} \psi \leftrightarrow A_b \psi \wedge Bel_a A_b \psi$

but under the assumption TE this time. The controlling agency of our modified toy scenario can then be described respectively as:

- $A_a(A_b \gamma \rightarrow \gamma)$: a indiscernibly throws a juggling ball at button 1, no matter what his beliefs are.

- $A_a(Bel_a A_b \gamma \rightarrow \gamma)$: a throws a juggling ball at button 1 no matter what, and also throws a juggling ball at button 2 whenever he believes that b tries to bring about γ .
- $A_a((A_b \gamma \wedge Bel_a A_b \gamma) \rightarrow \gamma)$: a throws a juggling ball at the push-button 1 whenever he believes that b tries to bring about γ .

Finally, essentially the same comments would be made about the resulting interpersonal controls, except that none of them would (necessarily) imply that γ holds.

Warneken & Tomasello’s experiments. In [45], Warneken & Tomasello describe four experiments of help behavior in prelinguistic or just-linguistic children. In one of them, the adult tries, or at least act as he tries, to put magazines into a cabinet. But the doors are closed and he bumps into it instead. The experiment⁸ shows that the infant helps the adult to achieve his task by opening the doors.

Say that *open* stands for the “cabinet is open”. The scenario can be formalized in the logic.

1. The subjective controlled agency: $Bel_{toddler} A_{adult} \text{open}$
2. The subjective controlling agency: $E_{toddler}(Bel_{toddler} A_{adult} \text{open} \rightarrow \text{open})$
3. Possibly: $A_{adult} \text{open}$

So, (1) the toddler believes that the adult tries to bring about that the cabinet is open, and (2) the toddler brings about that the cabinet is open when he believes that the adult tries to bring about that the cabinet is open. The subjective help captured by the experiment is then an interpersonal control

$$\text{GIC}(E_{toddler}, Bel_{toddler} A_{adult}, Bel_{toddler} A_{adult}, \text{open}).$$

It is a well-situated and effective interpersonal control. Also, it is a successful subjective help in the sense that $\vdash \text{GIC}(E_{toddler}, Bel_{toddler} A_{adult}, Bel_{toddler} A_{adult}, \text{open}) \rightarrow \text{open}$. (3) It is irrelevant whether the adult indeed tries to bring about that the cabinet is open, and the setting of the experiment does not allow us to conclude any way or the other. Hence, it is not an opportune event of assistance.

Acknowledgments

Emanuele Bottazzi was supported by the “Postdoc 2011” funding programme of the Autonomous Province of Trento, Italy (project “STACCO”). Nicolas Troquard was supported by a Marie Curie fellowship under grant PCOFUND-GA-2008-226070 (project “LASTS”).

⁸Captured in video http://www.eva.mpg.de/psycho/videos/children_cabinet.mpg.

References

- [1] Nuel Belnap and Michael Perloff. Seeing to it that: a canonical form for agentives. *Theoria*, 54(3):175–199, 1988.
- [2] Nuel Belnap, Michael Perloff, and Ming Xu. *Facing the Future (Agents and Choices in Our Indeterminist World)*. Oxford University Press, 2001.
- [3] Michael E. Bratman. Shared cooperative activity. *The Philosophical Review*, 101(2):327–341, 1992.
- [4] Roderick M. Chisholm and Dean W. Zimmerman. On the logic of intentional help. *Faith and Philosophy*, 13(3):402–404, 1996.
- [5] J. M. Darley and B. Latané. Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4):377–383, 1968.
- [6] Dag Elgesem. *Action theory and modal logic*. PhD thesis, Universitetet i Oslo, 1993.
- [7] Dag Elgesem. The modal logic of agency. *Nordic J. Philos. Logic*, 2(2), 1997.
- [8] John Martin Fischer. *The Metaphysics of Free Will*. Blackwell, Oxford, 1994.
- [9] John Martin Fischer. *Deep Control*. Oxford University Press, 2012.
- [10] Harry Frankfurt. Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66:829–39, 1969.
- [11] Harry Frankfurt. *The Importance of what We Care About*. Cambridge University Press, 1988.
- [12] Jonathan Gelati, Antonino Rotolo, Giovanni Sartor, and Guido Governatori. Normative autonomy and normative co-ordination: Declarative power, representation, and mandate. *Artificial Intelligence and Law*, 12:53–81, 2004.
- [13] Guido Governatori and Antonino Rotolo. On the axiomatisation of Elgesem’s logic of agency and ability. *Journal of Philosophical Logic*, 34(4):403–431, 2005.
- [14] Joseph Y. Halpern and Yoram Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artif. Intell.*, 54(2):319–379, 1992.
- [15] Vincent Hendricks and John Symons. Epistemic logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition, 2014.
- [16] Risto Hilpinen. On action and agency. In E. Ejerhed and S. Lindström, editors, *Logic, Action and Cognition: Essays in Philosophical Logic*, pages 3–27. Kluwer, Dordrecht, 1997.
- [17] Risto Hilpinen. Stig Kanger on deontic logic. In Ghita Holmström-Hintikka, Sten Lindström, and Rysiek Sliwinski, editors, *Collected papers of Stig Kanger with essays on his life and work: Vol. 2*. Kluwer, 2001.
- [18] J. Hornsby. *Actions*. Routledge and Keegan, London, 1980.
- [19] Jennifer Hornsby. Trying to act. In Timothy O’Connor and Constantine Sandis, editors, *A Companion to the Philosophy of Action*, pages 18–25. Wiley-Blackwell, 2010.
- [20] J. F. Horty. *Agency and Deontic Logic*. Oxford University Press, Oxford, 2001.
- [21] Stig Kanger. New foundations for ethical theory. In Risto Hilpinen, editor, *Deontic Logic: Introductory and Systematic Readings*, pages 36–58. Reidel, 1957/1971.
- [22] Stig Kanger. Law and logic. *Theoria*, 38:105–132, 1972.
- [23] Stig Kanger and Helle Kanger. Rights and Parliamentarism. *Theoria*, 32:85–115, 1966.

- [24] Shelley Dean Kilpatrick. Helping behaviour. In Roy F. Baumeister and Kathleen D. Vohs, editors, *Encyclopedia of Social Psychology*, volume 2, pages 420–424. SAGE Publications, Inc., 2007.
- [25] Lars Lindahl. *Position and Change – A Study in Law and Logic*. D. Reidel, 1977.
- [26] Emiliano Lorini and Andreas Herzig. A logic of intention and attempt. *Synthese*, 163:45–77, 2008.
- [27] Alia Martin and Kristina R. Olson. When kids know better: Paternalistic helping in 3-year-old children. *Developmental Psychology*, 4, 2013.
- [28] Alfred R. Mele. Intentional action: controversies, data, and core hypotheses. *Philosophical Psychology*, 16:325–340, 2003.
- [29] Susan Sauve Meyer. Aristotle on the voluntary. In Richard Kraut, editor, *he Blackwell Guide to Aristotle's Nicomachean Ethics*, pages 137–157. Blackwell, 2006.
- [30] Seumas Miller. *Social Action: a Teleological Account*. Cambridge University Press, Cambridge, 2002.
- [31] Elijah Milligram. Pluralism about action. In Timothy O'Connor and Constantine Sandis, editors, *A Companion to the Philosophy of Action*, pages 90–96. Blackwell, 2010.
- [32] Timothy O'Connor. Free will. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2014 edition, 2014.
- [33] Brian O'Shaughnessy. *The Will*. Cambridge University Press, 1980.
- [34] Philip L. Pearce and Paul R. Amato. A taxonomy of helping: A multidimensional scaling analysis. *Social Psychology Quarterly*, 43(4):363–371, 1980.
- [35] Ingmar Pörn. *Action Theory and Social Science: Some Formal Models*. Synthese Library 120. D. Reidel, Dordrecht, 1977.
- [36] Lambèr Royakkers. Combining deontic and action logics for collective agency. In Joost Breuker, Ronald Leenes, and Radboud Winkels, editors, *Legal Knowledge and Information Systems. Jurix 2000: The Thirteenth Annual Conference*, pages 135–146. IOS Press, 2000.
- [37] Felipe Santos and José Carmo. Indirect Action, Influence and Responsibility. In *Proc. of DEON'96*, pages 194–215. Springer-Verlag, 1996.
- [38] Felipe Santos, Andrew Jones, and José Carmo. Responsibility for Action in Organisations: a Formal Model. In G. Holmström-Hintikka and R. Tuomela, editors, *Contemporary Action Theory*, volume 1, pages 333–348. Kluwer, 1997.
- [39] John. R. Searle. Collective intentions and actions. In P. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, chapter Cambridge, MA. MIT Press, 1990.
- [40] Michael Smithson and Paul Amato. An Unstudied Region of Helping: An Extension of the Pearce-Amato Cognitive Taxonomy. *Social Psychology Quarterly*, 45(2):67–76, 1982.
- [41] Gerd Sommerhoff. The Abstract Characteristics of Living Systems. In F. E. Emery, editor, *Systems Thinking: Selected Readings*. Penguin, Harmondsworth, 1969.
- [42] Nicolas Troquard. Reasoning about coalitional agency and ability in the logics of “bringing-it-about”. *Autonomous Agents and Multi-Agent Systems*, 28(3):381–407, 2014.
- [43] Raimo Tuomela. *Cooperation - A Philosophical Study*. Springer, 2000.
- [44] Raimo Tuomela. *The Philosophy of Sociality: The Shared Point of View*. Oxford University Press, 2007.

- [45] Felix Warneken and Michael Tomasello. Altruistic helping in human infants and young chimpanzees. *Science*, 311(5765):1301–1303, 2006.
- [46] Felix Warneken and Michael Tomasello. The roots of human altruism. *British J. of Psychology*, 100:445–471, 2009.