2018, 4 (1)

ARGUMENTA

The Journal of the Italian Society for Analytic Philosophy

First published 2018 by the University of Sassari

© 2018 University of Sassari

Produced and designed for digital publication by the *Argumenta* Staff

All rights reserved. No part of this publication may be reproduced, stored or transmitted in any form or by any means without the prior permission in writing from *Argumenta*.

Editor
Masssimo Dell'Utri
(University of Sassari)

Editorial Board
Carla Bagnoli (University of Modena and Reggio Emilia), Francesca Boccuni (University San Raffaele, Milano), Stefano Caputo (University of Sassari), Massimiliano Carrara (University of Padova), Richard Davies (University of Bergamo), Ciro De Florio (Università Cattolica, Milano), Elisabetta Galeotti (University of Piemonte Orientale), Pier Luigi Lecis (University of Cagliari), Olimpia Giuliana Loddo (University of Cagliari), Giuseppe Lorini (University of Cagliari), Marcello Montibeller (University of Sassari), Giulia Piredda (IUSS-Pavia), Pietro Salis (University of Cagliari)

Argumenta is the official journal of the Italian Society for Analytic Philosophy (SIFA). It was founded in 2014 in response to a common demand for the creation of an Italian journal explicitly devoted to the publication of high quality research in analytic philosophy. From the beginning *Argumenta* was conceived as an international journal, and has benefitted from the cooperation of some of the most distinguished Italian and non-Italian scholars in all areas of analytic philosophy.

Contents

Editorial

The first issue of the fourth volume of *Argumenta* opens with a Special Issue, edited by Giuseppe Lorini and Wojciech Żełaniec, devoted to a topic that in recent years has been registering a growing, if sometimes rather critical, attention by a fair number of scholars: *constitutive rules*.

As the two editors make clear in their Introduction, constitutive rules highlight the fact that, over and above rules that prescribe and regulate, there are rules that create new forms of acts, facts, objects, events etc., where the facts in question are what John Searle has termed "institutional". The discovery of this new area of research has significantly influenced the philosophical analysis of norms and normativity.

This Special Issue comprises ten essays authored by some of the best minds currently pursuing this influential line of research, and focused on the task of carefully examining the diverse aspects of, and concepts that are germane to, the issue of the background of constitutive rules.

After the Special Issue, we present two articles: the former, entitled *Presentism and Causal Processes*, is an ingenious attempt to reconcile presentism and causation on the basis of a new approach to causation, while the latter, entitled *Indicative Conditionals as Strict Conditionals*, suggests that in a considerably wide class of cases the semantic and logical properties of conditionals can be elucidated by employing the resources of modal propositional logic. It is our conviction that these articles will greatly contribute to foster the discussion of their respective topics.

Argumenta 4, 1 (2018): 3-4
ISSN 2465-2334
© 2018 University of Sassari

The section of Book Reviews then rounds off the number once again. We are proud to offer readers three new thoughtful reviews of as many interesting books.

As always, all the articles appearing in *Argumenta* are freely accessible and freely downloadable. I heartily thank the Assistant Editors for their invaluable work on the preparation of this issue, and all the colleagues who have acted as referees.

Buona lettura!

Massimo Dell'Utri
Editor

Argumenta 4,1 (2018)
Special Issue

The Background of Constitutive Rules

Edited by

Giuseppe Lorini and Wojciech Żełaniec

The Journal of the Italian Society for Analytic Philosophy

Contents

The Background of Constitutive Rules
Introduction

Giuseppe Lorini and Wojciech Żełaniec***

* *University of Cagliari*

** *University of Gdańsk*

1. Two Revolutions in the Philosophy of Norms

In the twentieth century two revolutionary turns were accomplished in the philosophical research on norms and normativity: The first was the creation of a full-fledged deontic logic, a decade-long development which went through stages[1] before it culminated in the publication of the essay "Deontic Logic" by Georg Henrik von Wright in 1951. The second was the conceiving of the idea of a constitutive rule. The first turn opened a new field of study, to wit that concerning logical relations between norms (recall e.g. the question asked by von Wright (1951: 5) if there are any "logical truths peculiar to deontic concepts") and the logical structure of norms themselves; the second, by contrast, had consequences for the typological research of norms, the analysis of the functions that a norm can perform, and hence also the extension of normativity as such, that is, the inclusion in the extension of "normativity" of such phenomena as would not have been considered normative before the invention (or the discovery) of constitutive rules.

The present special issue of *Argumenta*, entitled "The Background of Constitutive Rules", is devoted to some aspects of that second revolution in the philosophy of norms, accomplished in the past century. It consisted, as we have briefly said, in putting forward the hypothesis that, beside the traditional roles of rules in regulating human behaviour by bestowing upon it deontic qualifications such as "obligatory", "prohibited", "permitted" or "facultative", some rules may well influence our life in a more radical manner, viz. changing the ontological and conceptual structure of our everyday world. Aside from the function of giving prescriptions and regulating (in the strict sense of this word, i.e. presupposing that that which is to be regulated already exists or at least is conceivable), a new normative function is revealed: the "constitutive" one. New ground is about to

[1] Regrettably, a true history of deontic logic has still to be written. However, in the words of Morscher (1975: 255), the book by Kalinowski (1972b) "*könnte ebensogut den Titel 'Geschichte der Normenlogik' tragen*" (could just as well have been entitled "A History of the Logic of Norms"), i.e. deontic logic, more or less. The German translation (Kalinowski 1972a) by Wolfgang Klein is in certain respects improved (by the translator).

Argumenta 4,1 (2018): 9-19
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.int

be broken: that issuing from the idea that rules not just regulate (types of) behaviour but are also apt to "do things with", to speak with J.L. Austin, and thus to have direct impact on reality, changing its ontological structure. In fact, by means of constitutive rules new forms of acts, facts, objects, events etc. are created, of which the relevant rules are conditions of possibility, not just in the ontological sense but, too, in the sense of sheer conceivability. Such acts etc., rendered ontologically and epistemologically possible (conceivable) by means of constitutive rules, came to be called "institutional" or "thetic" phenomena.[2]

2. A New Idea of Rule: Constitutive Rules

The first time ever (as far as we could establish) the idea of a constitutive rule makes its appearance in 1923 in Lemberg (Lwów), then Poland, where, in a Polish congress of philosophy, a young legal scholar and philosopher, Czesław Znamierowski, held a lecture with the title "Z nauki o normie postępowania" (From the research on the norm of conduct).[3] In this lecture Znamierowski distinguished between the traditional concept of a norm of conduct, which he believed is that of an "imperative" norm, and a new concept, or a new species of norms, which he called "construction norm". By way of a play-on-words one could be tempted to call that a distinction between "constriction norms" and "construction norms".[4] Three years later, in his book *Podstawowe pojęcia teorji prawa. Część pierwsza: Układ prawny i norma prawna* (Basic concepts for a theory of law. Part one: The legal system and legal norms) (1924: 72), he explained that a construction norm or a norm of construction (in Polish: *norma konstrukcyjna*) is one which, conferring new, conventional meaning on various types of acts and objects, constructs a new type of act.[5] As Znamierowski (1924: 103) chose to put it, "[a] norm of construction creates new possibilities of action, such as without it could not exist". For example, "without the rules of chess there would be no moves of the rook, the pawn or the queen, without the rules of contract bridge there would be no bids, tricks or passes".

According to Znamierowski, it is in virtue of construction norms (in his sense) that acts, actions, activities, as well as other kinds of psychophysical objects (Searle would call them by the Anscombean name of "brute" acts and so on) become what Znamierowski calls "thetic" (in Searle: institutional) acts, actions and so forth:

[2] The expressions "thetic acts/states of affairs" were introduced by Znamierowski (1924), while the expression "institutional facts" we owe to Searle (1964). At the origin of the Searlean concept of an "institutional fact" lies the essay "On Brute Facts" of 1958 by G.E.M. Anscombe.

[3] Cf. Znamierowski 1927.

[4] A likely source of inspiration for "construction norms", at least as far as terminology is concerned, was, for Znamierowski, the French legal scholar Léon Duguit, who as early as in his book *L'État, le droit objectif et la loi positive,* published in 1901, introduced the distinction between "*règles constructives"* and "*règles normatives*". On Znamierowski's social ontology see Lorini and Żełaniec 2013 and 2016.

[5] Here, the norm of construction (or a set of such norms) seems to perform an "ascriptive function": it ascribes a new meaning to things.

> Here, too, in virtue of a certain rule, or rather a whole system of rules, certain psychophysical actions acquire a new meaning constructed through norms. The chessmen, which are initially just pieces of lathed wood or bone, become "chessmen" that can be moved in this or that way (Znamierowski 1924: 67).

However, the philosopher who not only formulated his own version of the concept of constitutive rule but also made the very idea of such a rule well-known in various philosophical quarters was the American philosopher John R. Searle. In his essay "How to Derive 'Ought' from 'Is'" of 1964, he did not mention Znamierowski (who wrote in Polish) but, tracing back, perhaps not very accurately, his idea of the distinction between regulative and constitutive rules to Kant's distinction between regulative and constitutive principles,[6] he attributed it to John Rawls and his essay "Two Concepts of Rules" of 1955. Yet his distinction does have a striking similarity to that of Znamierowski mentioned above (i.e. the distinction between imperative norms and construction norms). One year after his essay of 1964 he writes:

> I distinguish between two sorts of rules: Some regulate antecedently existing forms of behaviour; for example, the rules of etiquette regulate interpersonal relationships, but these relationships exist independently of the rules of etiquette. Some rules on the other hand do not merely regulate but create or define new forms of behaviour. The rules of football, for example, do not merely regulate the game of football but as it were create the possibility of or define that activity (Searle 1965: 223).

Searle will call the former "regulative rules" and the latter "constitutive rules". As is evident from the excerpt above, these two different kinds of rules are distinct not just with regard to their function (regulating vs. creating) but also with regard to the relation that connects them to their respective objects. Regulative rules presuppose their objects as already existing, given; constitutive rules, by contrast, put these objects into existence. "Regulative rules regulate a pre-existing activity, an activity whose existence is logically independent of the rules". "Constitutive rules constitute [...] an activity the existence of which is logically dependent on the rules" (Searle 1965: 224).

In this very brief history of the research on constitutive rules a third scholar deserves a special mention: the Italian philosopher Amedeo Giovanni Conte. In the eighties of the twentieth century Conte developed a rich typology of constitutive rues, in which the subcategory of "eidetic-constitutive rules" was central.[7] Eidetic-constitutive rules are those that create the concept (*eidos*/εἶδος) of whatever they are constitutive of.[8] Conte characterises eidetic-constitutive rules not just in ontological terms (analogous to those employed by Znamierowski and Searle) but also in semiotic terms:

[6] As Conte has shown (2004a: 15; 2004b: 539ff.), Kant did come—if only in few little-known places in his work—somewhat closer to discovering the "Searlean" distinction than he could ever do by means of his distinction between constitutive and regulative principles.
[7] Cf. Żełaniec 2013: 75-93.
[8] Cf. Conte 1988 and Azzoni 1988.

> The eidetic-constitutive rules are those that determine the connotation of those terms that (in the formulation of the rules) designate the praxemes (the units of praxis) which are governed by the rules. For example the rules of chess are eidetic-constitutive because (and in the sense that) they determine the connotation of terms ('bishop', 'castling', 'check', …) which designate the praxemes (pieces, pragmemes,[9] game situations) of the game (Conte 1988: 252).

We have dwelt on these three key moments in the history of the research on constitutive rules, but clearly this history is much more complex and more minutely articulated.[10] In the twentieth century there were quite a few scholars, from various fields of research and various countries, who contributed in various ways, mostly independently from one another and having no knowledge of one another, to this new research topic, studying what has since Searle been called "constitutive rules" (in some cases eloquently denying their existence) under different names and in a variety of theoretical contexts. Amongst the most notable of these scholars were Carlos Alarcón Cabrera, J.R. Cameron, Gaetano Carcaterra, Stanisław Czepita, Jean-Louis Gardies, H.L.A. Hart, John Mabbott, Neil McCormick, Dolores Miller, Kazimierz Opałek, John Rawls, Alf Ross, David-Hillel Ruben, Dick W.P. Ruiter, Josef Schächter, Frederick Schauer, Hubert Schwyzer, D.S. Shwayder, Frank Snare, Ota Weinberger, Ludwig Wittgenstein, Jerzy Wróblewski, Zygmunt Ziembiński and others.

Beside these notable ones there are, too, some perhaps not yet so notable thinkers who have at some stage of their career done some serious work on constitutive rules, such as e.g. Frank Hindricks, or the editors of this volume. And, of course, there are still young scholars joining the constitutive rules research community. Some of them have contributed to this volume.

3. Is There Something Behind the Constitutive Rules?

The research on constitutive rules has from its beginnings onwards been inextricably connected with social ontology.[11] In social ontology, the fundamental role played by constitutive rules (referred to, again, by different names) in the "construction of social reality" (to quote Searle 1995) and its institutions with their normative implications was already acknowledged and the above-mentioned three pioneers of the research on constitutive rules: Znamierowski, Searle and Conte. Many various social facts and act(ivitie)s are institutionalised in one way or another, i.e. defined in terms of certain non-natural and man-bestowed properties and relations, and such institutionalising cannot be accomplished without constitutive rules. The case of the game of chess is paradigmatic:[12] as we have seen, it is impossible to play chess without the rules of chess. Generally, it is impossible to engage in any ever so implicitly institutionalised activity without the corresponding constitutive rules.

[9] Conte calls "pragmemes" the acts whose type is constituted by eidetic-constitutive rules. A castling in chess e.g. is a pragmeme.
[10] Cf. Conte 1986.
[11] Cf. Znamierowski 1924 and Searle 1995, 2010.
[12] However, it must not be thought that constitutive rules are relevant exclusively for (the study of) games and sports; this impression is sometimes generated by overworking that paradigmatic example.

However, some authors, such as Hubert Schwyzer and Dolores Miller, studying the role played by such rules in the "construction of social reality" (to quote Searle once again) have started exploring certain areas left unexamined by the three geniuses mentioned above. The areas that we mean extend into the "backstage" of constitutive rules. For instance, Schwyzer, in his essay "Practices and Rules" of 1969, started asking what is "behind" constitutive rules. Developing an idea of Wittgenstein,[13] Schwyzer sets up a thought-experiment involving the favourite example of all students of "rule-constituted activities", their *Drosophila melanogaster* as it were: chess. Let's imagine a society, he says, call it "Ruritanians",[14] in which chess is "played" not as a game but as a religious ritual. All chess rules known in our world are strictly adhered to, so it is correct to say that the activity in question is constituted by the same constitutive rules (the same eidetic-constitutive rules, in Conte's lexicon) as our chess (chess as a game) and yet there is no winning or losing in it, although it is still possible to checkmate one's "opponent", or to castle. But what kind of a strange rite is that? Schwyzer describes it thus:

> There is, he says, only one chess set for each community. Chess is enacted once every year by the priest of the community, for purposes of determining the will of the gods. If white mates black, the community and the crops will flourish; if black mates white, there will be trouble (Schwyzer 1969: 457).

This thought-experiment of Schwyzer is just one example amongst many—an example of what it should be the job of social ontology to investigate, but what social ontology has so far ignored: the (usually tacitly taken for granted) backdrop against which constitutive rules operate and without which they are not thinkable. For constitutive rules do not arise and work in an absolute vacuum: they are immersed in a conceptual and semantical atmosphere that conditions their normativity and their very meaningfulness. For example, while it is impossible to play chess (as we know it) without the rules of chess, it is likewise impossible—this is what Schwyzer's story, we believe, makes evident—to play chess (as we know it) if the constitutive rules are there and are known by most everybody, but the concept of game (a playful competitive activity) and the concept of victory do not exist—in a society like that chess as a game would remain unintelligible. Such concepts as "game" or "victory", however, are not for their part defined through constitutive rules—or at least, not through any such rules of the same order as those that define chess. Such concepts are, much rather, part of the *conceptual background* of these rules.

This conceptual background has usually been taken for granted. And this is no surprise: after all, *don't* we know what a game, and what the victory in a

[13] Wittgenstein (1953: I, § 200): "It is, of course, imaginable that two people belonging to a tribe unacquainted with games should sit at a chess-board and go through the moves of a game of chess; and even with all the appropriate mental accompaniments. And if *we* were to see it we should say they were playing chess."

[14] A name adopted from the titles of three novels by the English novelist and playwright Anthony Hope, *The Prisoner of Zenda* of 1894, *The Heart of Princess Osra* of 1896 and *Rupert of Hentzau* of 1898. It remains regrettably unexplored, to our knowledge, if there are any deeper connections, but the name, between Schwyzer's thought-experiment and Hope's political-*cum*-social fiction.

game is? … True, the concept of victory, or even that of a victory-in-a-game is very general, abstract, and for this reason rather "woolly", but when we familiarise ourselves with a new game, say "Class Struggle", and we learn that victory in that game consists in this or that, we immediately understand what is meant by that (in fact, we usually from the very beginning of getting familiar with a new game impatiently expect being told what the victory in this game consists in) and we do not make puzzled faces asking "but what do you mean by: 'that is the victory in this game'? In what way do you think this should enlighten me? It doesn't and it can't. I know what the victory-in-draughts or victory-in-Classic Slalom are, but these sports are so different from this game that I have no inkling as to how I should possibly transfer those notions on the game in hand. Draughts and Classic Slalom, besides, are very unlike each other too, so I can't see anything that their respective 'victories' (if 'victory' in abstraction from the game or sport had any meaning, and it probably has not) could conceivably have in common."

Yet, obvious as the concept of victory-in-a-game may appear be, it belongs precisely to that *background* of constitutive rules which for a long time was taken for granted. In this background various concepts are hidden which Dolores Miller (in her essay "Constitutive Rules and Essential Rules" of 1981) called "meta-institutional concepts". The meta-institutional concepts are concepts that go beyond (Greek: *metá*/μετά) the institutions of which they are conditions of possibility: they are presupposed by the constitutive rules of the relevant institution.[15] An example of a meta-institutional concept is, according to Miller, the concept of obligation. Miller holds that this concept is meta-institutional with respect to the concept of promise:

> *"Obligation"* is a concept which exists above and beyond promising—it is logically independent of promising, although promising may not be logically independent of it. Indeed, as Searle says, promising is an institutionalized form of obligation (Miller 1981: 188).

The presence of meta-institutional concepts reveals another important truth concerning the ontology of institutions (for Searle, "systems of constitutive rules"): institutions are not "categorially autonomous", they are no independent and complete, self-closed units.[16] They are immersed in a conceptual and possibly axiological atmosphere which conditions their possibilities of existence. Neither do they operate in a vacuum or *ex nihilo*; rather, they depend on that "atmosphere" for their constitutive force and their meaningfulness.

If this is so, then a philosopher of normativity is confronted with a new challenge: to examine that *background* of constitutive rules. In this connection, there arise questions like these:

What is there behind institutions conceived of as systems of constitutive rules?

What kinds of institutions are there accordingly?

[15] On the idea of "meta-institutional concept", see Lorini 2000, 2012, 2014 and Roversi 2010, 2014.

[16] Is it possible to construct a typology of institutions? For an attempt see Searle 2006.

Which are the conceptual structures that are presupposed by constitutive rules and are operative in the construction of social reality?

Beside constitutive rules with their background, are there any other necessary conditions of possibility of institutional facts?

and many others.

The ten essays that we submit to the reader in this special issue of *Argumenta* are devoted to these and similar questions, and attempt to shed some light on them.

4. The Ten Papers

This special issue devoted to "the background of constitutive rules" opens with the paper "Two Concepts of Constitutive Rules" by Jaap Hage. Hage, as distinct from many other students of constitutive rules, has no interest in the rules of games. His "rules" are first and foremost legal norms. He shows that there are, in fact, two kinds (not mutually exclusive) of constitutive rules. The first kind are rules that are part of a social practice (such as "legal persecution of theft", for instance) within which the rules can have different—including regulative—functions (such as defining theft, prohibiting it and bestowing on certain state functionaries the competence to punish the thief). The Searlean "counts-as" rules[17] (e.g. an action satisfying certain conditions is theft) are just a subset of them. The second kind are rules that create new facts by (symbolically) attaching them to certain facts already existing. Such rules need not (though they can) define practices; their main "point" lies in putting constraints on "possible worlds", i.e. on what can coexist with what (Hage deploys here a subtle and precise conceptual apparatus). Some of such rules (he calls them "dynamic") can create duties and/or obligations, which allows, Hage says, to redeem, even if with important modifications, the old Searlean "derivation" of Ought from Is.

In "The Occasions of Law (and the Occasions of Interpretation)", the second paper of this special issue, Frederick Schauer starts from Searle's insight "No remark without remarkableness" to study the significance of the Gricean concept of implicature in the legal worlds, and in particular on that which he calls "the occasions of law" and "the occasions of legal interpretation". Schauer argues that, just as an act of remarking presupposes some reason for the remark, a legislative act presupposes a reason for laying down a norm, and an act of interpretation presupposes a reason why the relevant rule cannot be (or: is better not) left without an interpretation. If this is right, he goes on to argue, then we have got a first glimpse of a "pre-legal world", silent, unspoken-of, a hidden background of legal norms, most often taken for granted. According to Schauer, this unspoken-of pre-legal world can be focussed on as soon as human behaviour breaches some unwritten and unspoken-of rules that make up that background.

The third paper, brief but conceptually rich is the one by John Searle, entitled just "Constitutive Rules". In it, Searle examines and reconstructs the concept of a constitutive rule which he first formulated in his essay "How to Derive 'Ought' from 'Is'" of 1964 and which he has ever since gone on revising in the light of his incrementally changing theory of the institutional reality, first formu-

[17] Cf. Searle 1969: 36.

lated in his book *The Construction of Social Reality* of 1995. Remarkably, this paper is the only essay in Searle's very large work which is uniquely dedicated to the concept of a constitutive rule. Aside from assembling various components of the definition of that concept, formulated elsewhere, Searle makes evident the enormous construction power of this conceptual tool due to its logical structure and its recursive applicability, as well as its function for the (comprehension of) human civilisation. (The reason why we have put "comprehension of" in parentheses here is that, first, as shown by Hage, constitutive rules are sometimes parts of social practices, and second, as argued for by Kaluziński, such rules must be—conceptually—familiar, and adhered, to by those who enact these practices.)

In "Searle on Normativity and Institutional Metaphysics", the fourth paper of the issue, William Butchard and Robert D'Amico address, them too, the problem of the Searlean derivation of Ought from Is. They attempt to show that the concept of constitutive rule, and more fundamentally that of a "constitutive connection" (i.e. such that fully accounts for, or "metaphysically grounds", that which it is constitutive of, while being no essential part thereof) between collective attitudes on one hand and obligations on the other, as long as these attitudes are normative in character, can be used to make that derivation appear plausible, without a threat of circularity. In an in-depth, extended polemic against David-Hillel Ruben, Butchard and D'Amico also make a broader case for the concept of "constitutive connections" as central to Searle's social ontology, frequently misunderstood, as they see it. In addition, they try to demonstrate that Searle's account of institutions, norms and normativity in general, though constructivist in a way, is nonetheless both naturalist and realist, in a sense which they carefully discuss and define.

In the fifth paper, entitled "Meta-meta-institutional Concepts? A Tale on Schwyzer and the Force of Technical Ends (Live from Ruritania)", Guglielmo Feis attempts to come to terms with the idea of a meta-institutional concept and level in the construction of social reality. Once we have admitted such levels, he asks, is it not likely that we'll sooner or later have to admit a further, meta-meta-institutional level, then perhaps, by the same logic, a meta-meta-meta-institutional level … and so on? In order to avoid infinite regress, Feis suggests, we had better adopt a different solution to the Schwyzerian challenge. Its core lies in seeing the "meaning" or "significance" of the given set of constitutive rules; one could call it the "point" of engaging in the practice they constitute. Chess as a game and chess as a hypothetical religious rite have the same rules and you can do without all those meta-levels, Feis says, if you attend to what people engage in playing chess *for*.

In his essay "Constitutive Rules: The Manifest Image and the Deep Image", the sixth paper of the issue, Maurizio Ferraris draws our attention to a recently discovered aspect of the background of social reality: it could be called "empirical background". Ferraris starts from two questions on constitutive rules: the question of the genesis of constitutive rules ("where do constitutive rules come from?") and the problem of their foundation ("what grounds constitutive rules?"). This is where the background of the constitutive rules comes in. Ferraris holds that "there is a layer of recordings and documents that grounds constitutive rules" and that these recordings and documents "are the cornerstone of the empirical backgrounds that warrants the production of social objects through constitutive rules". In other words, according to him, "constitutive rules require

a background of empirical conditions, which come from the role that recording playing in the production of these rules". In the light of this thesis Ferraris proposes to add a "deep image" of social reality, based on recordings and documents, to its "manifest image" as based on collective intentionality.

In the seventh paper, the one by Terry F. Godlove Jr. and entitled "Constitutive Rules, Normativity, and A Priori Truth", the author takes as his point of departure the seminal idea of a constitutive rule proposed by Searle in the sixties of the twentieth century, and submits to critical examination, along the same lines as Hubert Schwyzer (1969) (and more recently, as Giuseppe Lorini 2012, 2014), the conceptual background of constitutive rules, that is, the categorial background that makes possible the constitutive function of constitutive rules. However, as different from Schwyzer and Lorini, Godlove (adopting a generally Kantian point of view) does not advance any hypothesis of the existence of a "material background", i.e. a social, cultural and institutional contexts in which constitutive rules are embedded, but rather that of the existence of a "formal background", i.e. an *a priori* context of "higher-order rules that are constitutive, not of this or that form of behaviour, but of any form of cognitive engagement with the world".

In "What Does It Mean That Constitutive Rules Are in Force", the eighth essay in the issue, Bartosz Kaluziński, while he acknowledges, under a different name (he calls it a "deep convention"), the existence of what was to Lorini the meta-institutional level and to Feis the "meaning" of constitutive rules, he addresses an entirely different aspect of the issue. His question: "what does it mean that a set of constitutive rules are in force?" pertains to the very essence of what happens when human (or other rational) beings actually engage in an activity or behaviour defined by that set of rules. "Actually"—as distinct from "going through the motions" of that practice, or mimicking it or mocking it or … behaving in a way that has only an accidental similarity to that practice. Well, few students of constitutive rules have observed that such agents must know, accept and respect these rules, as well as intend to behave according to them. Kaluziński carefully fills this lacuna.

In the ninth paper, entitled "The Ludic Background of Constitutive Rules in Bernard Suits", Filip Kobiela focusses on the role which constitutive rules fill in Bernard H. Suits' regrettably yet-too-little known (at least among the students of constitutive rules) conception of what he calls "lusory attitude". Within Suits' account the meta-institutional level, or, alternatively, the meaning or the "point", of a given (constitutive rules-defined) game comes first, not last: "We play games in order to win (or in order to make our opponent win, sometimes), after all, don't we."[18] What comes second is the specific definition of winning in a given game and this can be checkmating the opponent, or scoring the highest number of goals, or arriving somewhere ahead of the others or … whatever else. Then, but only then, do the relevant constitutive rules come in, and their role is

[18] Actually, in the last FIFA World Cup in Russia there were strikingly many matches played, from a certain moment on, apparently with a view to *not* changing the score. This was not a case of the "to win or just for fun" alternative, as it was fun to nobody. German sports reporters have called this manner of playing "*Verwaltungsmodus*" or "paperwork processing mode". But then, FIFA World Cups are ruled by much more rules "than are dreamt of in your philosophy," Professor Suits!

to restrict—typically: very severely—the number of ways in which the event defined as winning (or, more accurately, the element that winning in the game in question is defined as) can be brought about.[19] Seen in this perspective, constitutive rules do no more contrast all that starkly with regulative rules (in Searle's sense). Suits suggests calling them constitutive-restrictive, constitutive-limitative, or constitutive-prohibitory rules. This is certainly a very interesting and novel way of seeing constitutive rules and their (in this case: "prelusory") background.

Finally, Corrado Roversi, in his paper "Constitutive Rules and the Internal Point of View", the tenth and last paper of this issue of *Argumenta*, maintains that in order to understand fully the life of an institution we must not restrict ourselves to constitutive rules, but we need to be acquainted also with the social purpose and meaning of the practice these constitutive rules are part of. In particular, he makes a distinction between five different perspectives on an institution, and he observes that only one of these perspectives is captured by constitutive rules. On this path, Roversi puts forward the idea of a three-dimensional social ontology consisting of a structure (framed by constitutive rules), a conceptual background, and an actual practice. It is only while examining these last two elements, he maintains, that we encounter concepts that are essential for the practice in question, although they are not reducible to concepts constructed by the relevant constitutive rules (i.e., castling and checkmating), namely: meta-institutional concepts (i.e., "victory" and "defeat") and para-institutional concepts (e.g. "filibuster" and "first-move advantage").

References

Anscombe, G.E.M. 1958, "On Brute Facts", *Analysis*, 18, 69-72.

Azzoni, G.M. 1988, *Il concetto di condizione nella tipologia delle regole*, Padova: CEDAM.

Conte, A.G. 1986, "Fenomeni di fenomeni", *Rivista Internazionale di Filosofia del Diritto,* 63, 29-57.

Conte, A.G. 1988, "Eidos. An Essay on Constitutive Rules", *Poznań Studies in the Philosophy of the Sciences and the Humanities*, 11, 251-57.

Conte, A.G. 2004a, "Kanon. Filosofia della regolarità", *Sociologia del Diritto,* 31, 3, 5-22.

Conte, A.G. 2004b, "Konstitutive Regel in Immanuel Kant", *Rivista Internazionale di Filosofia del Diritto,* 81, 537-48.

Duguit, L. 1901, *L'État, le droit objectif et la loi positive*, Paris: Fontemoing.

Kalinowski, G. 1972a, *Einführung in die Normenlogik*, transl. by W. Klein, Frankfurt am Main: Athenäum.

Kalinowski, G. 1972b, *La logique des normes*, Paris: Presses Universitaires de France.

Lorini, G. 2000, *Dimensioni giuridiche dell'istituzionale*, Padova: CEDAM.

[19] Otherwise, that event can be made to happen in an amazing variety of ways, some of which are quite remarkable; think e.g. of Rosie Ruiz and her "victory" in the 84th Boston Marathon in 1980.

Lorini, G. 2012, "Can Constitutive Rules Create a Practice?", *Praxis Filosófica*, 34, 139-48.

Lorini, G. 2014, "Meta-institutional Concepts. A New Category for Social Ontology", *Rivista di Estetica*, 54, 2, 127-39.

Lorini, G. and Żełaniec, W. 2013, "And Yet There Was Some: Czesław Znamierowski's Social Ontology", in Stancati, C., Givigliano, A., Fadda, E. and Cosenza, G. (eds.), *The Nature of Social Reality*, Newcastle: Cambridge Scholars Publishing, 182-96.

Lorini, G., and Żełaniec, W. 2016, "Czesław Znamierowski's Social Ontology and Its Phenomenological Roots", in Salice, A. and Schmid, H.B. (eds.), *The Phenomenological Approach to Social Reality. History, Concepts, Problems*, Dordrecht: Springer, 75-89.

Miller, D. 1981, "Constitutive Rules and Essential Rules", *Philosophical Studies*, 39, 2, 183-97.

Morscher, E. 1975, "Einführung in die Normenlogik: Zu einem Buch von Georges Kalinowski", *ARSP: Archiv für Rechts- und Sozialphilosophie / Archives for Philosophy of Law and Social Philosophy,* 61, 2, 255-62.

Rawls, J. 1955, "Two Concepts of Rules", *The Philosophical Review*, 64, 3-32.

Roversi, C. 2010, "Constitutive Rules in Context", *ARSP: Archiv für Rechts- und Sozialphilosophie / Archives for Philosophy of Law and Social Philosophy*, 96, 223-38.

Roversi, C. 2014, "Conceptualizing Institutions", *Phenomenology and the Cognitive Sciences*, 13, 1, 201-15.

Schwyzer, H. 1969, "Rules and Practices", *The Philosophical Review*, 78, 451-67.

Searle, J.R. 1964, "How to Derive 'Ought' from 'Is'", *The Philosophical Review*, 73, 43-58.

Searle, J.R. 1965, "What Is a Speech Act?", in M. Black (ed.), *Philosophy in America*, London: Allen and Unwin, 221-39.

Searle, J.R. 1969, *Speech Acts. An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.

Searle, J.R. 1995, *The Construction of the Social Reality*, London: Allen Lane.

Searle, J.R. 2006, "Social Ontology: Some Basic Principle", *Anthropological Theory*, 6, 1, 12-29.

Searle, J.R. 2010, *Making the Social World. The Structure of Human Civilization*, Oxford: Oxford University Press.

Wittgenstein, L. 1953, *Philosophical Investigations*, Oxford: Blackwell.

Wright, G.H. von 1951, "Deontic Logic", *Mind*, 60, 1-15.

Żełaniec, W. 2013, *Create to Rule. Studies on Constitutive Rules*, Milano: LED.

Znamierowski, Cz. 1924, *Podstawowe pojęcia teorji prawa. Część pierwsza: Układ prawny i norma prawna* (Basic concepts for a theory of law. Part one: The legal system and legal norms, in Polish), Poznań: Fiszer i Majewski.

Znamierowski, Cz. 1927, "Z nauki o normie postępowania" (From the research on the norm of conduct, in Polish), *Przegląd Filozoficzny*, 30, 348-49.

Two Concepts of Constitutive Rules

Jaap Hage

University of Maastricht

Abstract

In this article, it is argued that rules have two main functions, the practice-defining function and the constraining (fact-to-fact) function. These two functions are compatible. In their function as constraints, some rules are also indirectly regulative. In both of their functions, rules differ from the summaries (rules of thumb) that Rawls discussed and opposed to the constitutive (fact-to-fact) rules which make that some decisions are the right ones.

In his work, first on the philosophy of language and later on social ontology, Searle focused on one kind of constitutive rules: counts-as rules, which are constitutive in the sense that they attach new facts to the existence of "old" ones. In doing so, Searle created the scientific interest in constitutive rules which they deserve. However, because of his narrow focus on counts-as rules, Searle also created the impression that counts-as rules are all there is to constitutive rules. This impression is wrong, if only because it overlooks dynamic rules.

Keywords: Constitutive rules, Counts-as rules, Duties, Fact-to-fact rules, Obligations, Practice-defining rules, Regulative rules.

1. Introduction

Through the ground-breaking work of Searle, the notion of a constitutive rule received the attention it deserves. However, this very notion was used by Searle in two different senses. The one sense is that of a rule which partly defines a social practice by being part of it. I will call such constitutive rules "practice-defining rules". Constitutive rules in the other sense are rules which attach new facts to old ones. I will call such constitutive rules "fact-to-fact rules".[1] Perhaps because Searle did not distinguish the two notions clearly, he overlooked that counts-as rules, on which he focuses, are only one kind of fact-to-fact rules.

In this paper I will elucidate the distinction between the two kinds of constitutive rules, elaborate on fact-to-fact rules, and distinguish in that connection between static rules and dynamic rules. Moreover, I will also argue that there are two kinds of regulative rules—duty-imposing rules and rules that lead to obliga-

[1] In other work (Hage 2015b, 2016 and 2018) I use the term "fact-to-fact rules" for a subcategory of what is called "fact-to-fact rules" in this article.

Argumenta 4,1 (2018): 21-39
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.hag

tions—and that both kinds are also constitutive and only in an indirect sense regulative.

2. Rules that Define Social Practices

In *Speech Acts*, Searle distinguishes between regulative and constitutive rules by claiming that regulative rules regulate antecedently or independently existing forms of behaviour, while constitutive rules do not merely regulate but also define or create new forms of behaviour. An example of regulative rules would be the rules of etiquette, while the rules of football or chess would be examples of constitutive rules (Searle 1969: 33). Later, in *The Construction of Social Reality*, Searle repeats this characterization of constitutive rules, but the emphasis has shifted to the discussion of counts-as rules as a kind of constitutive rules (Searle 1985: 27-29 and 43-51).

Although Searle does not mention Rawls in *Speech Acts*, the idea of constitutive rules as rules that regulate a practice which itself depends on the rules that regulate it, seems to be inspired by Rawls' paper *Two concepts of rules* (Rawls 1955). In that paper Rawls distinguishes between the summary and the practice concept of rules.[2] Rawls' analysis of what he called the practice concept of rules, is very similar to Searle's notion of constitutive rules. Since Rawls' analysis is more detailed than what Searle wrote, I will try to elucidate this concept of constitutive rules through a brief discussion of Rawls' paper.

2.1 Two Levels of Justification

Rawls' central concern in his paper is the distinction between justification of a practice and justification of a particular action falling under a practice. More in particular Rawls tries to defend the utilitarian theory of punishment against the criticism that it would allow punishment of innocent persons.[3] The defence goes as follows. Utilitarianism contains a justification of the practice of punishing people who committed crimes, and this justification is consequentialist. The practice contains rules which specify under which circumstances a criminal can be punished. The justification of the punishment of a concrete person should be based on these rules. Let us assume that the practice of punishment is justified on utilitarian grounds and that it does not contain a rule that permits the punishment of innocent persons. In that case such punishment would not be justified, even though in this particular case punishment might be justified on utilitarian grounds. Attempts to justify punishment by pointing out the desirable consequences in a concrete case would be misguided, since (rule-)utilitarianism is not meant as a theory to justify the outcomes of concrete cases, but only as a theory for the justification of rules and the social practices to which they belong.

[2] This connection to Rawls' paper was also emphasized by Żełaniec (2013: 27-28). At the same place Żełaniec also mentions several other precursors who contributed to the theory of constitutive rules, whether or not under that name.

[3] I ignore here the complication that sanctions might not be called "punishment" (or even "sanctions") if inflicted upon innocent persons.

2.2 Two Concepts of Rules

Following up on his distinction between two levels of justification, Rawls distinguishes two concepts of rules. The first concept, which Rawls calls the "summary concept", takes rules to be rules of thumb which make it easier to determine what ought to be done according to some independent standard. This independent standard might, for instance, be the utilitarian one. Let us assume for the sake of argument that practically every case of punishing a thief would increase utility. This means that, according to the independent standard, almost all thieves should be punished. For courts that must take decisions about individual cases of punishment it is much easier to apply the rule that thieves should be punished, than to compute the costs and benefits of punishment in terms of utility for every case of theft that is brought before them. Epistemic efficiency suggests that judges apply the rule, rather than compute utility for individual cases. Given our assumption that the rule is merely an epistemic tool (a "rule of thumb"), the punishment of a concrete suspect is justified by utilitarian considerations. However, the judgment of the court is *epistemically* justified by the application of the rule.[4]

The second concept of rules, which Rawls calls the "practice concept", takes rules to be the ultimate justification of the decisions based upon them. For example, a thief can be punished because there is a rule to this effect, and not because punishment would maximize utility. The use of this rule, as part of the social practice of criminal law, may be justified on utilitarian grounds, but the justification of this general use should not be confused with the justification of a concrete instance of punishment.

To elucidate the distinction between the two concepts of rules, Rawls lists a number of characteristics of rules on the summary concept and on the practice concept, and opposes them. For my present purposes the relation between the rules and the cases to which they are applied is particularly interesting. According to Rawls, on the summary concept, the decisions made on particular cases are logically prior to rules. He means that it is determined whether a person can be punished independently of the rule on punishment. This determination results, for instance, from application of the utilitarian standard. Moreover, the rule of thumb is correct or incorrect depending on whether its outcomes match the outcomes of the utilitarian standard. On the practice concept of rules, however, the rules are logically prior to the proper decisions in particular cases, because these decisions depend on the rules that must be applied. This means that the correctness of the rules cannot be tested against the results they produce in a concrete case.

When I wrote that in the practice concept of rules, the rules are logically prior to the proper decisions in particular cases, I (intentionally) misrepresented Rawls a bit. Rawls, actually, wrote that "the rules of practices are logically prior to particular cases" (Rawls 1999: 36). By that formulation he meant to say that some kinds of behaviour are only possible because there are rules regulating this behaviour. An example would be the rule of article 310 of the Dutch Criminal Code, which, amongst other things, defines which kinds of behaviour count as theft. The punishment of thieves (those who committed theft) is only possible

[4] This use of the expression "epistemically justified" stems from the present author, not from Rawls.

thanks to this rule which defines what counts as theft. This kind of logical priority does not follow from the distinction between the summary concept of rules and the practice concept. That some rules constitute, rather than predict, the outcomes of cases does not mean that these rules also define particular kinds of events. Therefore I reformulated Rawls' characterization of the logical priority of rules on the practice concept to a version in which the logical priority does follow. However, the fact that the logical priority of the kinds that Rawls had in mind does not follow from the distinction between the two concepts of rules, does not mean that this kind of logical priority is uninteresting. On the contrary, it is this kind of logical priority which Searle had most likely in mind when he used it to found the distinction between regulative and constitutive rules.

2.3 The Ontological Priority of Constitutive Rules

We must distinguish between the priority of rules over concrete decisions for which Rawls made a case and which is based on the distinction between the summary concept and the practice concept of rules and the priority which Searle uses to argue for the distinction between regulative and constitutive rules. I will call the former priority "logical" and the latter priority "ontological". The ontological priority of constitutive rules boils down to this: these constitutive rules are necessary for the existence of the phenomena with which these rules deal. For instance, rules which make thieves punishable can only make sense if they also define theft. I use on purpose the rather vague verb "to deal", because we will see that there are several reasons why it cannot well be said that constitutive rules also guide the phenomena which they constitute.[5]

Let us first look how this ontological priority functions in the classical example of playing chess.[6] Some forms of behaviour, such as moving sculptured pieces of wood over a board with a pattern of light and dark squares, count as playing chess partly because of the existence of the rules of chess. These rules make playing chess possible and are in that sense constitutive for chess and therefore they are called *constitutive* rules. However, these rules do not prescribe behaviour and if "regulating" is interpreted as determining what is and what is not allowed, the constitutive rules of chess do not regulate the game. They determine what counts as a valid move in chess, they determine what the players should strive for in order to play the game seriously, they define certain situations such as check-mate and stale-mate, and they determine when a game is finished and who the winner is. The rules of chess "deal with" playing chess, but they are not prescriptive or permissive.[7] If Searle distinguishes regulative rules

[5] To anticipate the argument that will follow: one reason is that constitutive rules typically do not prescribe behaviour at all (this subsection), not even in a wide sense, while the other reason is that if rules 'prescribe' behaviour in a wide sense, strictly speaking they do not prescribe at all, but merely constitute obligations or duties (section 5.1).

[6] The example is classical, not only because it was used by Searle (1969: 33-34), but also because it was used to more or less the same purpose by Alf Ross (1959: 11-17). Ross was in turn inspired by Kelsen (1934/1992: 10-11), who did not use the chess example, but wrote about norms as schemes of interpretation, a kind of constitutive rules *avant la lettre*.

[7] So I disagree with Żełaniec when he writes (Żełaniec 2013: 103) that, once it has been established that a particular item is a bishop, 'we must put that item under the *obligation* to move diagonally only' (italics in the original). Żełaniec raises this point to argue that the rule that a bishop can only move diagonally is not constitutive for what counts as a

from constitutive ones only by means of the practice which they deal with, he seems to overlook that regulative rules are by definition deontic (prescriptive or permissive), while constitutive rules very often are not. Therefore it is less than felicitous to say that constitutive rules "regulate" the practice which they constitute.

It might be objected that the rules of chess also regulate the game by prohibiting some moves ("you are not allowed to castle when your king is in check") and allowing others ("on its first move a pawn is allowed to go two fields forward"). This objection does not cut ice, though. Characteristic of a prescription is that violation is normally possible. In case of the rules of chess, violation leads to invalid moves which have no influence on the progress of the game. These rules, which define the game, cannot be violated because a violation would require that an illegal move be made in the game, while "moves" that violate the rules do not count as moves in the game.[8] Therefore it is not possible to violate the rules that define how pieces move in a game of chess, which is a sure sign that these rules are not prescriptive.

Criminal law is a social practice which is in part constituted by mandatory (prescriptive) rules. This practice consists of cooperating organizations (police, courts, prisons) the functioning of which is governed by rules. Moreover, the rules that belong to the practice define kinds of behaviour (e.g. theft), prohibit these kinds of behaviour (the mandatory rules), and empower courts to impose sanctions in case these prohibitions were violated. As this example illustrates, rules are only part of the practice, where the practice does not only consist of rules, but also of organizations. Moreover, the rules which belong to the practice, and in that way partially constitute the practice, fulfil rather different roles. They include counts-as rules, such as the rule that defines what counts as theft[9], but also mandatory rules, such as the rule that prohibits theft, and power-conferring rules[10], such as the rule that makes it possible to sanction criminals.

All these kinds of rules—counts-as, mandatory, and power-conferring—contribute to the existence of the practice and are in that sense constitutive. So it is not only counts-as rules which are constitutive rules in the sense of practice-defining rules.

It is worthwhile to consider how counts-as rules, mandatory rules and power-conferring rules operate together to partly constitute the social practice of criminal law. If we take the prohibition of theft, we see a mandatory rule which regulates behaviour which both does and does not exist prior to the existence of the prohibition. Would it be possible to have thefts without a rule that prohibits theft? It seems so. It is true that theft in the sense of criminal law cannot exist without a rule defining what counts as theft, but this counts-as rule does not

bishop in a particular instance of the game. Perhaps Żełaniec is right on that score, but this rule is constitutive for what counts as a valid move for a bishop, and does not impose any obligations on bishops in chess, or on chess players.

[8] If the pieces are made of wood, the only thing that moves in case of an illegal move, is the piece of wood, and perhaps also the chess-piece (the piece of wood in the meaning it has in chess). The game did not move (progress), however.

[9] I do not want to suggest that theft cannot exist without such a counts-as rule. However, counts-as rules are essential for the existence of theft *in the sense of criminal law*.

[10] I only use the expression 'power-conferring rule' here because it is established. In *Foundations and Building Blocks of Law* (Hage 2018: 200) I argue that, strictly speaking, there are no power-conferring rules, but that there are competence-conferring rules, which are not the same thing under a different name.

prohibit theft. For the existence of theft in the sense of the criminal law, the existence of a counts-as rule is required, but not the existence of a mandatory rule. To this extent, the rule that prohibits theft does not constitute the behaviour that it regulates. However, this rule makes it possible to violate the criminal law, and such a violation is required for the competence of a court to sanction the criminal.[11] Although it is possible to commit theft even if theft is not prohibited, and the prohibition can therefore be said to regulate antecedently existing behaviour, the counts-as rule that defines theft only makes sense in the context of a social practice which also contains the prohibition and the power-conferring rules. It is only possible to commit theft in the sense of criminal law if the social practice of criminal law exists, and all the different kinds of rules that go into criminal law are conditions for the existence of this practice, and therefore indirectly also for the existence of concrete cases of theft. In this sense, the prohibition of theft *is* ontologically prior to concrete cases of theft.

2.4 Conclusion on Practice-Defining Rules

Starting from Searle's distinction between regulative and constitutive rules and Rawls' distinction between the summary and the practice concept of rules, we found that social practices depend for their existence on rules. These rules are ontologically prior to facts that exist in the context of these social practices, and are in this sense constitutive for both the practices and the facts that exist in their context. I call these constitutive rules practice-defining rules and emphasize that practice-defining rules may be of many different kinds, including counts-as rules, mandatory rules and power-conferring rules.

In subsection 3.3 we will encounter fact-to-fact rules, which are constitutive in a different sense, because they function as constraints on possible worlds. However, first I will provide a setting for these rules by going into some detail concerning directions of fit, and constraints on possible worlds (subsections 3.1 and 3.2).

3. Rules as Constraints on Possible Worlds[12]

Often the notion of a rule is connected to the guidance of behaviour: rules indicate what we should do. Typical examples are the rule that impose the duty on home-owners to clean away the snow on the pavement in front of their houses and the rule that prohibits building in a zone that the municipality reserved for environmental purposes. However, there are also rules whose primary function does not seem to be to guide behaviour. Examples would be the rule that gives the municipality council the power to make a parking regulation for the city and the rule that makes persons against which a serious suspicion exists that they committed a crime into criminal suspects in the sense of the Code for Criminal Procedure.

If it is not the primary function of all rules to guide behaviour, the question is whether there is a function that all rules share. This question can be answered

[11] For the sake of argument I ignore the complication that the crime needs to be proven if the competence is to arise.

[12] The argument in this section is a re-working of the argument of Hage 2015a. See also Hage 2018: 57-70.

affirmatively: all rules have in common that they attach the presence of facts to the presence of other facts. A proper understanding of this common characteristic requires that we pay some attention to what will be called the "world-to-word direction of fit" of rules.

3.1 Directions of Fit

Perhaps the best way to introduce the distinction between directions of fit is by means of an example of Anscombe's (Anscombe 1976: 56). Suppose that Elisabeth makes a shopping list, which she uses in the supermarket to put items in her trolley. A detective follows her and makes a list of everything that Elisabeth puts into her trolley. After Elisabeth and the detective are finished, the list of the detective will be the same as Elisabeth's shopping list. However, the lists had different functions. If Elisabeth used the list correctly, she placed exactly those items into her trolley that are indicated on the list. Her behaviour is to be adapted to what is on her list. In the case of the detective it is just the other way round; the list should reflect Elisabeth's shopping behaviour. The two different functions of the lists with regard to Elisabeth's behaviour represent the two different directions of fit that we are looking for.

The two items involved in Anscombe's example are a linguistic one—the list of items—and the world—the collection of all facts. The directions-of-fit distinction can also be applied to other items than linguistic ones, but let us focus on the linguistic case first. The relation between language and the world goes in two directions. If the linguistic entities, in particular descriptive sentences, are to be adapted to the world, as when the detective writes down which groceries are in the trolley, the fashionable expression is "word-to-world direction of fit". If the world is to be adapted to the linguistic entities, as when Elisabeth puts those items in her trolley that are mentioned in her shopping list, the fashionable expression is "world-to-word direction of fit" (Searle 1979: 3-4).

For the world-to-word direction of fit we must distinguish between three kinds. For all three kinds it holds that somehow the facts in the world are adapted, in order to "fit" what is expressed by the words. One case is when the words function as a *directive*, as when James shouts "Carol, stop!" when he fears that his young daughter Carol will cross the busy street. This order aims at making Carol stop, and if the order is successful in the sense of "efficacious", Carol will stop and the facts in the world fit the content of the order. In this case the relation between the utterance of the order (the performance of the speech act) and the facts in the world is causal. I will therefore write about the "causal world-to-word direction of fit".

A second form of the world-to-word direction of fit manifests itself in constitutive speech acts. Constitutive speech acts are speech acts performed with the intention to bring about a particular change through the operation of a rule or convention. They differ from directives which operate by means of a causal, rather than a rule-based, connection. Examples of constitutive speech acts are the baptism of a ship ("I hereby baptize you the Princess Victoria"), making a promise ("I promise you to be back before 8 o'clock this evening"), granting a power ("You can consider every promise made by Michal on my behalf as a promise made by me"), and the issuing of a command (as distinguished from an order). In all these cases the facts in the world have come to fit the content of the speech act, but in contrast to the operation of directives, the result is brought about by a

rule or convention, and not through a causal connection. In connection with constitutive speech acts we will speak of the "constitutive world-to-word direction of fit".

The result of a successful command, such as "I hereby forbid you to cross the street" directed by James to his daughter Carol, is that a duty enters into existence.[13] In this case it is the duty for Carol not to cross the street. If such a command is successful, the facts in the world come to match the content of the speech act and Carol has from that moment on the duty not to cross the street. In this case the relation between the speech act and the facts in the world is constitutive by nature; the performance of the speech act constitutes the duty. This is a case of the constitutive world-to-word direction of fit, and I mention it explicitly to emphasize the difference between orders, conceived as a kind of directives, and commands, conceived as constitutive speech acts. The terminological distinction between orders and commands is stipulative: that is how I will use these words here. However, the difference between directives, based on a causal connection, and the creation of duties by means of constitutive speech acts and based on the operation of rules, does not depend on this terminological convention.

The third kind of world-to-word direction of fit concerns the effects of rules. Take for example the conceptual rule (the meaning postulate) that the word "rectangle" denotes quadrilaterals with four orthogonal angles. Given this rule, if something is a rectangle, it must be a quadrilateral with four orthogonal angles. This "must" depends on the conceptual rule that defines the relation between being a rectangle and being a quadrilateral. The facts in the world adapt themselves to the rule—the quadrilateral now also is a rectangle—and that is what is meant by the world-to-word direction of fit of rules, or—as we will see in section 3.3—more in general the world-to-word direction of fit of constraints.

The idea that rules have a special form of the world-to-word direction of fit deserves an elaborate explanation, but that requires more background on possible worlds and necessity. That is the topic we will therefore turn to now.

3.2 Possible Worlds

We are all familiar with the distinction between what the facts actually are and what the facts might have been. It happens to be snowing, but the sun might just as well have been shining. In Syria there is a war, but there might have been peace.

Logicians use possible worlds-terminology to deal with this distinction between what the facts actually are and what they might have been.[14] They say, for instance, that in the actual world it is snowing, but that in some other possible world the sun is shining. Intuitively, a possible world is an exhaustive set of facts which makes some descriptive sentences true and others false. The actual world is one of the many worlds that are possible and in the actual world it is snowing. However, in some other possible world, the sun is shining. That is another way of saying that although actually it is snowing, the sun might have been shining.

[13] It is *not* that Carol does not cross the street. That would be a causal effect of Carol's belief that she has the duty not to cross the street.

[14] To keep the exposition relatively simple, the following accounts of possible worlds semantics and of propositional logic are not very precise.

What is necessary is the case in all possible worlds, while what is impossible is not the case in any possible world. What is contingent is the case in some, but not in all possible worlds. In all possible worlds, nine is the square of three. Moreover, in all possible worlds, if Luis is either in Madrid or in London, and he is not in London, then he is in Madrid. However, in some possible worlds, Luis is in Madrid, while in some other possible worlds he is in London.

3.3 Constraints on Possible Worlds

The intuitive idea about the connection between possible worlds and necessity is quite simple, but there turns out to be a complication: what is the difference between a possible world and an impossible one? Rather than trying to answer this question in abstract, it is better to study an example in which the idea of possible worlds has turned out to be fruitful. This example is propositional logic.

As its name indicates, propositional logic deals with propositions, the meanings of descriptive sentences. Some propositions are basic or atomic; they do not exhibit any internal structure. Other propositions are compound or molecular. They contain a logical operator and—most of the times—more than one basic proposition. Examples, with the logical operators italicized, are "Jane loves Mary *and* Mary loves Jane", "*If* Siobhan talks to Joachim, *then* Joachim listens", and "It does *not* rain".

An important difference between elementary and compound propositions is that the truth-values (true or false) of the former are completely independent from one another, while the truth-values of the compound propositions are completely determined by the truth-values of the elementary propositions that are part of them. To be more concrete, the truth-values of "Jane loves Mary" and of "Mary loves Jane" are, logically speaking, independent from each other, while the truth-value of "Jane loves Mary *and* Mary loves Jane" completely depends on the truth-values of the former two propositions. The latter proposition is true if and only if the two former propositions are both true.

If two worlds agree in their truth-values for "Jane loves Mary" and "Mary loves Jane", and if they do not agree in their truth-values for "Jane loves Mary *and* Mary loves Jane", at least one of these worlds cannot be logically possible. Logically possible worlds are constrained in such a way that the truth-values of compound propositions in these worlds are determined in a particular way—depending on the form of the compound sentences and the logical operators that occur in them—by the truth-values of the elementary propositions that constitute these compound propositions.

Constraints on logically possible worlds can be characterized in terms of constraints on the facts that are elements of these possible worlds. In a logically possible world some compound states of affairs must obtain if some elementary states of affairs obtain, and the other way round. For example, the compound state of affairs "Jane loves Mary *and* Mary loves Jane" must obtain in a logically possible world that contains the elementary states of affairs "Jane loves Mary" and "Mary loves Jane". More in general, worlds that are possible according to propositional logic satisfy the constraints of propositional logic. A world that does not satisfy these constraints does not count as logically possible.

Constraints on possible worlds make that some things are necessarily the case, while other things cannot be the case. Given the constraints of propositional logic, it must be the case that either it rains or it does not rain. Given

these same constraints it cannot be the case that both it rains and it does not rain. However, these constraints make it possible that both it rains and it is snowing.[15] Constraints on possible worlds do not only support necessity and possibility judgements, but also conditional judgements ("conditionals"). For an arbitrary possible world we cannot know whether it snows, but we do know that if it snows, then the state of affairs "It snows or the sun shines" obtains. The reason why constraints support modal judgements—judgements about what is (im)possible and what is possible—and conditionals, is that constraints have the world-to-word direction of fit. They determine which worlds are possible and in that way impose themselves on possible worlds. A world in which it rains can only be logically possible if either it rains or the sun shines. Possible worlds are adapted to the constraints to which they are subjected. Of course this adaptation is not something that happens in time. It is a conceptual matter; the world only counts as logically possible if it satisfies the constraints of logic.

The constraints of propositional logic are useful to illustrate the operation of constraints on possible worlds, because they are well-defined and—at least to the readers who know propositional logic—quite familiar. However, other constraints than the logical ones are also important. Physical constraints, for example. A world in which sound travels faster in vacuum than light is not physically possible, at least not according to our present knowledge of physics. The same holds for a world in which metals do not expand when heated, or a world in which the total amount of energy fluctuates. The constraints of physics limit the facts that can physically go together.

Just as logical constraints underlie logical necessity judgments, physical constraints underlie physical necessity (and possibility) judgments. We have already seen some examples of judgements that particular facts cannot go together (impossibility judgements). Examples of physical necessity judgements are that a piece of metal necessarily expands when heated, that light must travel in a right line through vacuum if not subject to gravitational forces, and that an unsupported body with gravitational mass must drop when in the neighbourhood of the Earth. Physical constraints also support physical conditionals and counterfactual judgements: "If this piece of metal were heated, it would expand" and "Had this stone not been supported, it would have dropped".

What holds for physical constraints also holds for conceptual constraints (meaning postulates; semantic conventions). The convention that defines bachelors as unmarried males underlies the judgment that necessarily all bachelors are unmarried, while the convention that defines what a book is underlies the judgment that it is possible to read books. Semantic conventions also support conditionals such as "If this is a skate board, it is a vehicle in the sense of the Traffic Act", and counterfactuals such as "If this vehicle had a motor, it would have counted as a car".

The last two examples concern semantic rules that could also have been legal rules. In general, conceptual constraints depend for their existence on their being adopted by a language-using community. They illustrate that constraints can depend for their existence on being adopted by human beings.

[15] Perhaps snow and sunshine cannot go together but that would not be a matter of logic, but of what is physically possible. Some things are logically possible, but physically impossible.

3.3 Rules as Constraints on Possible Worlds

Why are rules a kind of constraints? Because they behave like other constraints. In a world in which a rule exists, the rule imposes itself on the facts of that world with the world-to-word direction of fit that other constraints also have. So if some possible world contains the rule that thieves are punishable, then in that world thieves are punishable. In that world, it is not merely a contingent matter of fact that thieves are punishable, but a necessary one, because being a thief makes one punishable. Moreover, the rule also supports conditional and counterfactual judgements: if Jane had been a thief, she would have been punishable.

Rules have a lot in common with more traditional constraints such as the logical and physical ones, but they also have a characteristic that is not shared by all other constraints: rules only apply locally. The laws of one country are for example different from the laws of another country. The necessity of rule-based judgements is therefore merely local necessity: in some African countries practicing homosexuals count as criminals, while that is not the case (anymore) in European countries. This is different for logical and physical laws, which seem to have a universal scope of application.[16] The scope of rules is not only limited in space, but also in time. Many rules can be created or repealed and in that sense they differ from more traditional constraints which somehow seem outside the scope of human manipulation. When the rule that thieves are punishable is introduced, suddenly all thieves become punishable. And when the rule is repealed again, the possibility to punish thieves disappears with the rule.

As a consequence of these differences, there can be some logically and physically possible worlds in which a particular rule exists, and other possible worlds in which the same rule does not exist. In a sense it might be said that logical and physical constraints create necessities that are themselves necessary, while rules create contingent necessities. For this reason, rules will be categorized as "soft constraints", as opposed to the hard constraints that seemingly do not depend for their existence on human decision making or social practices.[17]

3.4 Factual and Descriptive Counterparts of Rules

The rule that thieves are punishable makes it impossible that thieves are not punishable.[18] Or, to state the same thing affirmatively, the rule necessitates that thieves are punishable. If some rule—or, more in general, a constraint—exists, this means that some general descriptive sentence must be true.

[16] This difference should not be overestimated, however. The geometrical law that the three corners of a triangle add up to 180 degrees only holds for relatively small triangles and (which may be the same issue) for triangles in a flat plane. See also the discussion of the scope of physical laws in Toulmin 1953: 69 and 78.

[17] 'Seemingly', because it is not clear in which respect physical laws are more than mere regularities, with their aspect of necessity added through social practice. Similarly, with the proliferation of logical systems it becomes increasingly clear that logical necessity may be more mind-dependent than used to be assumed. This is an important theme, closely related to the nature of constraints and the way they differ from mere regularities, but it deserves a more extensive treatment than it can be given here.

[18] For our present purposes we will not take into consideration the possibility of exceptions to rules.

Such sentences are open generalizations which deal with potentially infinitely many items. The open generalizations that describe the effects of rules typically have the same formulation as the rule the effects of which they describe, and they are true because that rule exists. If a rule has the formulation "Thieves are liable to be incarcerated for a maximum of five years", the effect of this rule can be described by saying that thieves are liable to be incarcerated for maximum five years. The sentences that describe the consequences of a rule describe facts that will be called the "factual counterpart" of the rule.

The sentences themselves may be called the "descriptive counterparts" of rules. Where rules impose themselves on the world by way of their world-to-word direction of fit, but are not true or false, the descriptive counterparts of rules are descriptive sentences with the word-to-world direction of fit, which are true or false, usually depending on the existence of the rules of which they are the counterpart.

Since the facts which make these sentences true obtain because of rule, descriptive counterparts of existing rules must be true. The facts which are expressed by these sentences and which exist because of rules can be designated as the factual counterparts of these rules.

The differences between rules, their factual counterparts and their descriptive counterparts, may be confusing, but the three can be kept apart if one realizes that rules have the world-to-word direction of fit (they constrain the facts in the world), that descriptive counterparts have the word-to-world direction of fit (they describe the facts in the world), and that facts are just elements of the world and have no direction of fit.

4. Fact-to-Fact Rules

All rules have the world-to-word direction of fit: they impose themselves on the worlds (places and times) in which they exist by attaching facts to other facts. That is why they are also called "fact-to-fact rules". The connected facts may be simultaneous, or they may succeed each other in time.

4.1 Dynamic Rules

The latter is the case with dynamic rules: they create new facts, or modify or take away existing facts as a consequence of the occurrence of an event. One example is that Jane promised Richard to give him €100, with the consequence that from the moment of the promise on Jane has the moral obligation to give Richard €100. Another example is that Eloise was appointed chair of the French Parliament, with the consequence that from the starting point of the chair's new term on, Eloise will be the chair of the French Parliament.

These examples also illustrate that the world-to-word direction of fit of constitutive speech acts is based on the world-to-word direction of fit of dynamic rules. Because of the operation of the dynamic rule that making a promise creates an obligation, promising is a constitutive speech act with the world-to-word direction of fit that it creates an obligation. The distinction between the world-to-word direction of fit of constitutive speech acts and the world-to-word direction of fit of constraints and more in particular dynamic rules may be useful for expository purposes, but they are two sides of the same coin.

4.2 Static Rules

Whereas dynamic rules govern the succession of facts in time, static rules govern the co-existence of facts. They make (constitute) that one kind of fact goes together with some other kind of fact, where the latter fact depends (supervenes) on the former. The relation between the kinds of facts is timeless, in the negative sense that the one kind of fact is not the occurrence of an event after which the second kind of fact comes into existence.

Legal examples of static rules are that:

1. the owner of a good is permitted to use this good;
2. home-owners must (have the duty to) keep the pavement in front of their houses clean;
3. the Mayor of a municipality has the power to issue emergency regulations for that municipality;
4. the King of the Belgians is the commander in chief of the Belgian army.

4.3 Counts-as Rules

One kind of fact-to-fact rules has become particularly well-known: the *counts-as rules*. Counts-as rules have the structure: Individuals of type 1 count as individuals of type 2. These "individuals" may be human beings, as in the rule that the parents of a minor count as the minor's legal representatives, or the rule that the king of the Belgians is the commander in chief of the Belgian army. Often, however, the "individuals" that count as another kind of individual are events. For instance, under particular circumstances, causing a car accident counts as committing a tort, or offering money to another person counts—given "suitable" circumstances—as attempting to bribe an official. Frequently counts-as rules are conditional, meaning that individuals of type 1 only count as individuals of type 2 if certain conditions are satisfied. An example from Dutch law (art. 3:84 of the Civil Code) would be the rule that the delivery of a good counts as the transfer of that good if the person who made the delivery was competent to transfer and if there was a valid title for the transfer.

4.4 Practice-Defining and Fact-to-Fact Rules

All rules are fact-to-fact rules, and if there are practice-defining rules, they must be fact-to-fact rules at the same time. When we look at examples of practice-defining rules, for examples the rules that constitute the social practice of criminal law, we find confirmation of this claim. There are counts-as rules which define the crimes, there are duty-imposing rules, which forbid acts that would constitute a crime, and there are power-conferring rules, granting courts the power to impose sanctions upon convicted criminals. These rules, in combination with each other and with organizations which apply them, together constitute the practice of criminal law, and they are all fact-to-fact rules at the same time. So there is no opposition between practice-defining rules and fact-to-fact rules: sometimes fact-to-fact rules (partly) define a social practice and then they are also practice-defining rules. There is no opposition, but only a contrast between different functions which rules can fulfil, and there is no reason why a rule cannot fulfil both functions at the same time.

5. Regulative Rules

5.1 Duties and Obligations

The English language does not make a strict distinction between duties and obligations, but in the sphere of normativity there is an important distinction and the words "obligation" and "duty" are quite suitable to mark this distinction.[19] Obligations are relations between typically two concrete persons, where the one person, the debtor, is obligated to do something, or to refrain from doing something, while the other person, the creditor, has a claim on the debtor that he perform his obligation. Obligations are the result of an event that took place, typically a promise, a contract, or a tort, to which the obligation was attached by a dynamic rule. The content of the obligation is partially defined by this event. Examples of obligations are the obligation that Jane undertook towards Richard by promising him that she would pay him €100, and the obligation of Carol towards Nina to compensate the damage that Carol caused to Nina by bumping into Nina's car.

Duties, on the contrary, do not involve other persons than the duty-holder, although the content of the duty may mention a person. For example, Heinrich may have the duty not to hurt his dog Bonzo and also the duty not to hurt his brother Franz when they have a quarrel. Neither Bonzo nor Franz has a claim against Heinrich not to be hurt, although Franz may receive a claim against his brother for damages if his brother breaches his duty not to hurt Franz. While obligations are attached to events by a dynamic rule, duties are typically[20] attached by a static rule to the possession of a certain status. For example, a static rule attaches the duty to clean the pavement in front of a house to the status of being the home-owner, while another static rule attaches the duty to turn on the car lights when it gets dark to the status of being a car-driver. The most general duties are attached to the status of being human, or to the status of a legal subject.[21]

We find that a person can in at least two ways become obligated to do something: via a dynamic rule and via a static rule. In the first case the dynamic rule creates an obligation; in the second case the static rule creates a duty. Both the dynamic rule and the static rule are in themselves not deontic: they do not obligate but they create, respectively an obligation and a duty. It may seem that rules that create obligations or duties are nevertheless normative because of their normative impact. To some extent this is indeed the case, but they are not so very different from other dynamic and static rules. A dynamic rule which creates an obligation is not very different from a dynamic rule which makes somebody into the chair of the French parliament. The only difference between them is the kind of fact which they attach to an event. An obligation-creating rule makes that some obligation exists, while the rule which makes somebody into the parliament chair makes that somebody is the parliament chair. A static rule which attaches the duty to clean the pavement to the status of home-owner is not very

[19] The suitability has to do with the use of the word 'obligation' in the civil law tradition (Zimmerman 1996: 1-10), which matches the use of the word that will be proposed here.
[20] Typically, but not always: a duty may be the result of a command.
[21] Notice that the status to which a static rule attaches duties (or other consequences) may, but need not, itself be the result of a fact-to-fact rule.

different from the rule that attaches the power to dispose of the house to this same status.

The impact of this observation that regulative rules, that is obligation-creating and duty-imposing rules, are not very different from dynamic, respectively static rules which do not have normative consequences can be seen clearly in the following reinterpretation of Searle's "derivation" of Ought from Is.

5.2 Searle's "Derivation" of Ought from Is

It is often assumed that Is and Ought represent different ontological spheres, and that therefore it is not possible to derive ought-conclusions from merely is-premises. That would give regulative rules a special ontological status, because they would provide the deontic content for duties and obligations. In a famous article Searle attempted to show that this alleged gap between Is and Ought can be bridged through the speech act of promising. This would undermine the special role of regulative rules, because they would not be necessary anymore for the existence of obligations. Searle's conclusion was correct, be it that his argument can be improved. The improved version which is presented in this section provides us with additional insight into the nature of fact-to-fact rules and is for that reason included here.[22]

Searle produced his famous argument by means of which he tried to show that it is possible to deduce Ought from Is alone in an article from 1964 (Searle 1964). In this argument an actual derivation of an ought-conclusion from is-premises takes the central role. The derivation went as follows:

1. Jones uttered the words "I hereby promise to pay you, Smith, five dollars".
2. Jones promised to pay Smith five dollars.
3. Jones placed himself under (undertook) an obligation to pay Smith five dollars.
4. Jones is under an obligation to pay Smith five dollars.
5. Jones ought to pay Smith five dollars.

Searle argued that the relation between a statement in this list and its successor is either an entailment or at least not contingent, and moreover that the relation could, where needed, be made into an entailment by the addition of a premise which was neither an evaluative statement, nor a moral principle, nor anything of the sort.

Searle's argument is less than convincing, but nevertheless correct in its underlying idea. It is less than convincing because of the presupposition needed to get from (2) to (5), which is presumably something like:

For any x and any A, if x promised to do A, then x ought to do A.

The problem is that this presupposition is a sentence which expresses, among other things, also an ought. Searle tries to tackle this by the claim that the sentences which make the transition between the sentences of the argument deductively valid are analytic (literally: "tautologies"), but even if Searle is cor-

[22] The argument was presented earlier, in a more elaborate form and in a different setting, in Hage 2011 and Hage 2013. See also Hage 2018: 97-102.

rect in this claim, the issue remains that analytic ought-sentences are still ought-sentences. Therefore, Searle did not succeed in deriving Ought from Is alone.[23]

Nevertheless, Searle's argument that it is possible to derive Ought from Is has an underlying idea which is correct. This underlying idea is that some events lead to *new* obligations and therefore new oughts, and that the presence of these events itself does not depend on pre-existing obligations or oughts. In the following, the correctness of this idea is illustrated by using contracts as example, instead of promises.

5.3 Contracts and their Consequences

Contracts are a means by which legal subjects can change the legal positions of themselves or other persons. Contracts can only operate in a setting of rules, which define, amongst other things, how contracts can be made, which persons are competent to and have the capacity to contract and which legal consequences are connected to the successful creation of a contract. These latter rules are the dynamic rules which play such a crucial in Searle's argument.

The only point of making promises is to undertake obligations. Although it is possible to undertake obligations by means of contracts as well, contracts can be also used for other purposes that cannot be achieved by promises. It is for instance possible to appoint by means of a contract an arbiter who is empowered to decide over conflicts that might arise in connection with the execution of (the rest of) the contract. In a contract, the parties create or repeal facts, to the extent that they are empowered by dynamic rules to do so. Although contracts do not necessarily lead to obligations, they often do. And when contracts lead to obligations, they almost always also lead to the oughts that follow from these obligations.[24] In this way, contracts seem to bridge the gap between Is and Ought.

It may be objected that contracts only seemingly do so because the obligation to do what was contracted is based on the regulative rule that contracts ought to be complied with. The contract itself would on this view be nothing more than a specification for a concrete situation of what the general obligation to perform contracts, laid down in the regulative rule, implies. However, it is questionable whether the rule that contracts ought to be complied with really exists. The point of contracts is more general than merely that contracts facilitate the intentional creation of obligations. Their point is that the facts established by means of the contract hold between the contract parties.

Underlying contracts is not the rule that contracts ought to be performed, but the dynamic rule that the facts which a contract aims to bring about, actually come into existence. The rule that what parties agreed to holds between the parties, does itself not impose any obligations. If obligations result from most contracts, this is because by means of most contracts the contract parties create obligations between themselves. Notice the emphasis on "create". Before the contract was concluded, the obligations were not there yet; they are the result of the contract. The presumed rule that one ought to obey one's contracts is superfluous. If the contract does not create obligations, but aims for instance at cancel-

[23] See Hare 1964.

[24] It is possible that somebody has the obligation to do something, for instance because he contracted to do so, while at the same he legally ought to refrain from doing so. More on the relation between obligations and what ought to be done in Hage 2018: 148-52.

ling existing obligations, there is nothing to obey. If the contract does create obligations, the rule that one ought to comply with these obligations would effectively be that one ought to do what one is under an obligation to do. That would be an almost analytical rule. So, apart from this analytical rule according to which an obligation typically leads to an ought, there is no role for the rule that contracts ought to be obeyed. The obligation to do what one contracted to do does not, therefore, derive from such a rule. This obligation is *in a concrete case* created by means of the contract and it is a new obligation that did not yet exist before the contract, not even in the more abstract form of an obligation to comply with one's contracts.[25]

Searle's derivation of Ought from Is rests on two steps. One step is the logical representation of constitution. Given a setting of rules, some events lead to the coming into existence of new facts, and these facts may include the existence of obligations. This mechanism is the main step in Searle's argument, and the interesting thing about it is that it has almost nothing to do with oughts or obligations. It is a mechanism by means of which all kinds of facts in social reality are generated by means of rules. That some of these facts involve the existence of obligations is almost a coincidence. The second step is the "derivation" of an ought from the existence of an obligation. This is also an important step, but it is not important for the point that Searle—perhaps unknowingly—made, namely that through the application of rules—in his example dynamic rules—new facts can come into existence, and that these new facts may very well involve the existence of obligations.

The relevance of this point can hardly be overestimated, because it illustrates the distinction between two kind of normativity. On the one hand there is the kind of normativity that is also called "deontic". It has to do with what agents have the duty or the obligation (not) to do. On the other hand there is the normativity of facts that are the result of rule-application, rule-based facts, which are sometimes also called "institutional facts".[26] It is tempting to treat these two kinds of normativity as one and the same thing, a temptation which is only strengthened if one assumes that all oughts must be rule-based. However, it is possible to create new facts by means of contracts and other juridical acts and these new facts can, but certainly need not, involve obligations. The conventional world-to-word direction of fit of constraints has nothing to do with duties, obligations and what agents ought to do. These latter entities have at best—that is: if they are complied with—a causal world-to-word fit.

5.4 Conclusion on Regulative Rules

Regulative rules create either duties or obligations. They *create* them, which means that before the duty, respectively the obligation, was created, there was nothing mandatory. Not even the regulative rule itself was mandatory, because it merely creates, constitutes new facts. Rules that create duties typically men-

[25] Notice, by the way, that this "obligation" to comply with one's contracts cannot even be an obligation. It holds in general, is not based on a specific event, and has no creditor. In other words, if it had existed, it would have been a duty.

[26] See Anscombe 1976, MacCormick 1986, and Searle 2010. It has always escaped me why these facts should be called "institutional", since only in a limited number of cases, social institutions are involved in the existence of rule-based facts.

tion the duties which they create, and as a consequence it may *seem* that they involve the duties themselves. For example the rule that home-owners have the duty to clean the snow in front of their houses mentions the duty to clean the snow. However, if there are no home-owners, there is no duty, not even if the rule exists. Therefore the mere existence of a duty-imposing rule does not involve the existence of a duty.

In the case of a rule that creates an obligation this is even more pronounced. Some rules that lead to an obligation at least mention the obligation. This is the case with the rule that creates the obligation to pay damages in case of a tort. However, the rule that the facts to which contract parties agreed hold between the parties, does not even mention an obligation. If this rule leads to an obligation, this is because the contract parties agreed on the creation of an obligation.

Rules that impose duties or create obligations—in short: regulative rules—are merely constitutive, and if they lead to an ought, a *new* ought, it is only the created ought that is regulative. The rule itself is merely constitutive. This means that all rules are constitutive and that strictly speaking no rule is regulative. This need not stop us from calling duty-imposing rules and *some*[27] obligation-creating rules regulative, as long as we are aware that they are only indirectly regulative, namely through the duties and obligations that they constitute.

6. Final Conclusion

In this article I have argued that rules have two main functions, the practice-defining one and the constraining (fact-to-fact) function. These two functions are compatible. In their function as constraints, some rules are also indirectly regulative. In both of their functions, rules differ from the summaries (rules of thumb) that Rawls discussed and opposed to the constitutive (fact-to-fact) rules which *make* that some decisions are the right ones.

In his work, first on the philosophy of language and later on social ontology, Searle focused on one kind of constitutive rules: counts-as rules, which are constitutive in the sense that they attach new facts to the existence of "old" ones. In doing so, Searle created the scientific interest in constitutive rules which they deserve. However, because of his narrow focus on counts-as rules, Searle also created the impression that counts-as rules are all there is to constitutive rules. This impression is wrong, if only because it overlooks dynamic rules. A possible consequence is that Searle underestimated the relevance of his derivation of Ought from Is in his early paper.

References

Anscombe, G.E.M. 1976, *Intention*, 2nd ed., Oxford: Basil Blackwell.

Hage, J.C. 2005, *Studies in Legal Logic*, Dordrecht: Springer.

[27] Only some, because only some dynamic rules have the creation of obligations as a legal consequence. Other dynamic rules have more abstractly defined consequences, which *may* involve obligations.

Hage, J.C. 2007, *De wondere wereld van het recht* [The Wondrous World of the Law], Inaugural Address, University of Maastricht.

Hage, J.C. 2011, "Legal Transactions and the Legal Ought", in Stelmach, J. and Brożek, B. (eds.), *The Normativity of Law*, Kraków: Copernicus Press, 167-90.

Hage, J.C. 2013, "Juridical Acts and the Gap between Is and Ought", *Netherlands Journal of Legal Philosophy*, 42, 50-66.

Hage, J.C. 2015a, "Separating Rules from Normativity", in Araszkiewicz, M., Paweł Banaś, Gizbert-Studnicki, T. and Pleszka, K. (eds.), *Problems of Normativity, Rules and Rule-Following*, Cham: Springer, 13-30.

Hage, J.C. 2015b, "The (Onto)logical Structure of Law: A Conceptual Toolkit for legislators", in Araszkiewicz, M. and Pleszka, K. (eds.), *Logic in the Theory and Practice of Legislation*, Cham: Springer, 3-48.

Hage, J.C. 2016, "Facts, Values and Norms", in Taekema, S., van Klink, B. and Been, W. (eds.), *Facts and Norms in Law*, Cheltenham: Edward Elgar Publishing, 23-52.

Hage, J.C. 2018, *Foundations and Building Blocks of Law*, Den Haag: Eleven International Publishing.

Hare, R.M. 1964, "The Promising Game", *Revue Internationale de Philosophie*, 70, 1964, 398-412. Also in Foot, P. (ed.), *Theories of Ethics*, Oxford: Oxford University Press, 1967, 115-27.

Kelsen, H. 1992, *Introduction to the Problems of Legal Theory*, transl. of the first edition of the *Reine Rechtslehre* (1934), Litschewsky Paulson, B. and Paulson, S. (eds.), Oxford: Clarendon Press.

MacCormick, N. 1986, "Law as Institutional Fact", in MacCormick, N. and Weinberger, O. (eds.), *An Institutional Theory of Law*, Dordrecht: Reidel, 49-76.

Rawls, J. 1955, "Two Concepts of Rules", *Philosophical Review*, 64, 3-32. Also in his *Collected Papers*, Cambridge, MA: Harvard University Press, 1999, 20-46.

Rawls, J. 1999, *Collected Papers*, Freeman, S. (ed.), Cambridge, MA: Harvard University Press.

Ross, A. 1959, *On Law and Justice*, Berkeley: University of California Press.

Searle, J.R. 1964, "How to Derive 'Ought' from 'Is'", *Philosophical Review*, 73, 43-58. Also in. Foot, P. (ed.), *Theories of Ethics*, Oxford: Oxford University Press, 1967, 101-14.

Searle, J.R. 1969, *Speech Acts*, Cambridge: Cambridge University Press.

Searle, J.R. 1979, *Expression and Meaning. Studies in the Theory of Speech Acts*, Cambridge: Cambridge University Press.

Searle, J.R. 1995, *The Construction of Social Reality*, New York: The Free Press.

Searle, J.R. 2010, *Making the Social World*, Oxford: Oxford University Press.

Toulmin, S. 1953, *The Philosophy of Science. An Introduction*, New York: Harper Row.

Żełaniec, W. 2013, *Create to Rule. Studies on Constitutive Rules*, Milano: Edizioni Universitarie di Lettere Economia Diritto.

Zimmerman, R. 1996, *The Law of Obligations. Roman Foundations of the Civilian Tradition*, Oxford: Oxford University Press.

The Occasions of Law (and the Occasions of Interpretation)

Frederick Schauer

University of Virginia

Abstract

When John Searle observed that there is "no remark without remarkableness," he made a point about the pragmatics of conversation that is importantly applicable to legal interpretation. Just as the act of remarking, according to Searle, presupposes some reason for the remark, so too does the act of legal interpretation presuppose a reason to interpret. This paper explores this phenomenon, and identifies the distinct occasions that call for an act of interpretation.

Keywords: Law, Legal interpretation, Hard cases, Easy cases.

1. Introduction

"No remark without remarkableness" is one of the philosopher John Searle's most profound insights.[1] It may be true, for example, that Professor X is not drunk, but if I *say* that Professor X is not drunk then something else is going on, and my statement has a more complex meaning. More specifically, by saying—remarking—that Professor X is not drunk, I am implying that there is a reason for saying that he is not drunk. Searle's important point is that the most plausible reason for saying that Professor X is not drunk is that there is something remarkable—something worth remarking about—about his not being drunk. In some contexts this remarkableness might stem from the way in which a person might be different from other people. If some person were two meters in height, for example, my asserting that "You are two meters tall" would be worth asserting precisely because most people are not that tall, and thus there is something remarkable about this person's height. But in other contexts the remarkableness is a remarkableness within some individual and not across individuals, and thus when I remark that Professor X is not drunk, the more plausible implicature—philosophical terminology for the conversational implication[2]—is that Professor X is drunk on other occasions. And so although the literal meaning of "Profes-

[1] Searle 1969: 144-45.
[2] See Grice 1989, Davis 2014.

Argumenta 4,1 (2018): 41-50
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.sch

sor X is not drunk" is that Professor X is not drunk, the implication of saying it—of remarking on it—is that he is drunk on other occasions.

Searle's basic and profoundly important idea, therefore, is that an assertion—a remark—presupposes a reason for offering the assertion, and that the most typical presupposition lying behind a true assertion is the plausibility of its negation. We assert something as true only when its not being true is plausible, or conceivable. That plausibility might come, as I have just noted, from the existence of the negation for other agents, or it might come from the existence of the negation for the same agent (or other object of the assertion), or it might come from the existence of some person actually asserting the negation. If one person says that the earth is round, the plausibility, in context, of the negation may come from a setting in which some other person has (sincerely) asserted that the earth is flat.

2. The Occasions for Rules

If Searle's point is sound, and I believe that it is, then much the same idea applies to rules, another topic about which Searle's writings over the years have been extremely and properly influential.[3] Rules, or more precisely regulative rules in Searle's typology, do not, of course, assert. They prescribe, and thus they say what ought to be the case and not what is the case.[4] Prescriptions may order, or command, or request, or suggest, or recommend, but in some way they indicate what it is that the speaker wishes to take place. But then, in the same spirit as "no remark without remarkableness", we can again observe that a prescription presupposes the (empirical) plausibility of the behavior that the rule prohibits or that the prescription seeks to have occur.

A few examples will illustrate the point. Some years ago there were signs on the Massachusetts Turnpike, a high speed limit access motorway extending from the state's western border to the Atlantic Ocean on the east, that instructed motorists that they were not permitted to drive backwards (against the flow of traffc) on the Turnpike if they happened to miss their intended off-ramp. For most drivers in most places such a warning would seem absurdly superfluous, because going backwards on a limited access motorway is so dangerous that no one in their right mind would think of doing it. But the very existence of the warning in Massachusetts tells us something about Massachusetts drivers. After all, there must be a reason why Massachusetts sees a need to warn against behavior that in most other places would seem simply beyond comprehension. And thus the fact that driving in reverse on the Massachusetts Turnpike is prohibited, and the fact that the prohibition is the subject of a specific warning, tells us that the behavior is genuinely conceivable for Massachusetts drivers, even if the behavior is not for more normal drivers in more normal places.

Consider to the same effect the rules of practice for the Supreme Court of the state of Wyoming. Among the rules for the behavior of lawyers arguing cases before the Court are rules governing the time that is allocated to each party for an argument, the requirements for written filings (briefs) with respect to length, form of citation, and so on. But most interesting is the rule (technically a

[3] Especially Searle 1969.

[4] For my own analysis of rules, and on rules as prescriptive generalizations, see Schauer 1991.

guideline, but presumably an enforceable one) instructing Wyoming lawyers as to how they should address the judges of the court during an oral hearing, including one instruction telling lawyers that they should not refer to the judges as "you guys".[5]

"You guys"??!! One might reasonably have thought it inconceivable that a lawyer during the formal context of a legal argument would address the members of the court as "You guys", but plainly such an assumption would be mistaken, at least in Wyoming. Were it not for the plausibility of the behavior that most of us think unthinkable, Wyoming would not have thought it necessary to prohibit it. Wyoming's rule thus signals the plausibility of behavior whose existence would not otherwise have crossed our minds.

For a third example, consider the Third Amendment to the Constitution of the United States, a part of the Bill of Rights, which was ratified and added to the original 1787 Constitution in 1791.[6] In many respect, of course, many of the world's constitutions are quite similar to each other, protecting similar rights and otherwise doing at least some similar things.[7] But the United States Constitution, unique among the constitutions of the world, prohibits the "quarter[ing] of troops in private homes". It does, of course, seem like a very bad idea for the state to require that homeowners convert their homes into barracks for the housing of soldiers, whether for the short or the long term, but it is such a bad idea that no modern state has thought it necessary to guard against it. In the United States of 1791, however, the recollections of the British doing just that during the period prior to (and during) the American Revolution were sufficiently salient to justify inclusion in a Bill of Rights of a provision that would now seem to most constitutional drafters in most countries as superfluous.

Finally, and more briefly, I note an article in a relatively recent issue of a scuba diving magazine, the title of the article being "Don't Pet the Sharks".[8] I am a scuba diver myself, and it has never occurred to me to pet a shark. But the title of the article, even without more, has informed me that others obviously feel differently, and the prohibition has provided information about the empirical plausiblility of that which I had previously believed implausible. And so too with the warnings at the British Midland Airways counter at Heathrow Airport in London, reminding people that it is against the law to assault an airline or airport employee. Again, we can assume that the warning arises out of the genuine likelihood that people will indeed assault airline counter employees. I suspect, but do not know, that one does not find similar warnings in countries with smaller amounts of violence against service employees. And thus if positive remarks—assertions—presuppose remarkableness, then negative prohibitions presuppose a certain kind of non-remarkablenes. They presuppose the plausibility of the behavior that the prohibition prohibits.

3. The Occasions of Law

All of these examples, and the deeper point that they are designed to illustrate, can tell us a great deal about what I call "the occasions of law". Not only is law

[5] Wyoming 2014: 229-35.
[6] U.S. Const. amend. III.
[7] See Ginsburg, Elkins and Simmons 2013: 61-95.
[8] *Sport Diver* 2010: 3-4.

not inevitable, but specific instances of law—whether they be prohibitions, requirements, permissions, or something else—exist against a background of non-law. Law in the broadest sense is an exception, and a great deal of human behavior and human interaction takes place without the intervention of law.[9] It is only when something goes amiss that law is called upon to remedy the gaps and the problems that exist in our pre-legal existence.

Not only is law an exception in this sense, but it is also not the only exception. In one of his later essays, Hans Kelsen described law as a "specific social technique".[10] And although it may not be especially noteworthy to describe law as a social technique, Kelsen's description of the social technique as "specific" is telling. What Kelsen understood, and properly so, was that there are various techniques of governance, of coordination, of control, and of much else available to, figuratively, a society's institutional designer, with law being merely one of those various and multiple techniques.

Thus, insofar as there exists a social desire for behavioral change, or, as most of my examples were designed to illustrate, a social desire to protect certain existing behavioral patterns and social norms against outliers, the designers of social institutions have a number of alternatives for achieving those goals. Consider again the example of the prohibition on the petting of sharks. To the best of my knowledge, petting sharks is not currently unlawful in most of the waters of the world, including the waters that are part of the territorial borders of various nations. But now suppose, realistically, that there develops a concern about such behavior, either because of its environmental consequences to the sharks, or because of the danger to the people who pet them. In the face of such concern, one possible response would be to make it illegal to touch or approach a shark. But the important point is that this is not the only possible response. Another alternatve might be a series of speeches by respected public figures warning of the undesirability of petting sharks. Still another might be the construction of underseas barriers separating the sharks from divers. And still another might be an educational campaign in schools (as we now see with respect to environmental and climate change issues) designed to inculcate an anti-shark-petting norm from an early age.

The example is somewhat silly and unrealistic, but is designed to illustrate Kelsen's basic point—that law exists not as a universal antidote to all of society's problems or as a universal approach to all of society's goals, but instead as a particular technique arguably suited for some problems and goals but not for others. Determining when law is appropriate and when it is not will require specification of just what we mean by law in this context, and then delving more deeply into the fit (or absence of fit) between law and various social problems and social goals. This is not the occasion, in part for reasons of space, to explore either of these issues. Nevertheless, Kelsen's basic idea is instructive, because it highlights from a different direction the non-universality and the non-ubiquity of law. Lawyers, of course, are apt to overstate the importance of their own enter-

[9] Obviously a great deal here turns on the definition of "law", but for present purposes I understand law simply as the state-connected institution for the creation and enforcement of regulative and constitutive social rules, an institution including enacted and published rules, courts, judges, lawyers, and police officers. The conception of law with which I work here is thus relevantly similar to that developed in Hart 2012.

[10] Kelsen 1941: 75-95 and Kelsen 1957: 231-56.

prise in the grand scheme of things, but that pathology is hardly unique to lawyers. I have little doubt that dentists, architects, artists, and barbers all do the same thing, although it does not appear that dentists and barbers are as hegemonic as lawyers about the pervasiveness and importance of their enterprise. Still, it is worthwhile being cautious when listening to lawyers talk about the importance of law. This is not to say that law is unimportant. But it is to say that many other things are important as well, and that law occupies only one corner of our larger social existence.

4. The Problems of Making Law with Hard Cases

Although the examples with which I commenced this essay were designed to highlight the way in which legal rules and legal control more generally arise in response to specific needs, specific threats, and specific instances of the lack of legal control, it is worth pointing out the way in which legal responses to perceived needs for law may often be misguided, and in a particular way. Specifically, it is often the case that the need for a legal rule arises from a particular incident. Perhaps, for example, that only one lawyer had ever referred to the judges of the Supreme Court of Wyoming as "you guys", but, as is often the case, a single glaring event may be the impetus for legal change or the creation of a legal rule.

Insofar as this is true, however, and I admit it may be more frequent in common law systems where specific controversies may provide the platform for judicial law-making, it does suggest that the process of law-making may be unsystematic in rarely noticed ways. Specifically, let us begin with the proposition that any rule, and therefore any legal rule, covers multiple instances. Rules are general, and that is part of what makes them rules.[11] And because rules are by their nature general, it is the task of the law-maker to imagine the field of instances that some rule will cover. To make a rule prohibiting vehicles in the park, for example,[12] the rule-maker must imagine the field of vehicles, or, more specifically, the field of vehicles that might, absent the rule, be driven into the park.

So far so good, but now a new dynamic comes into play. Led by Daniel Kahneman and the late Amos Tversky, the research agenda known as heuristcs and biases studies those irrationalities of human decision-making that often impede sound and rational judgment.[13] And although Kahneman and Tversky and their followers have identified a large number of such irrationalities, the one that is especially relevant here is what has come to be known as the *availability heuristic*.[14] The basic idea is simple—we imagine that that which is closest to us, or easiest to see—that which is most available—as being more representative of some larger set or larger class than it actually is. If we go to a party and meet ten people, three of whom as doctors, we may think that doctors comprise a larger percentage of the population than is in reality the case. If there happened to be

[11] See Schauer 1991. See also, and relevantly to the argument in the text here, Schauer, 1995: 633-59.

[12] The example, ubiquitous in jurisorudential writings, comes originally from Hart 1958: 593-629. For (too) extensive analysis, see Schauer 2008: 1109-1135.

[13] See, for example, Tversky and Kahneman 1974, 1981 and 1986. For overviews, see Kahneman, Slovic and Tversky 1982, Plous 1993.

[14] Tversky and Kahneman 1973. See also Reyes, Thompson and Bower 1980.

an earthquake yesterday, we are similarly likely to believe that earthquakes are more common than they actually are.

And so too with the specific events that prompt the making of legal rules.[15] When faced with a highly salient event seemingly in need of a legal or rule-based response, the availability heuristic warns us that we or the rulemaker may think the particular event is more representative than it actually is, and predictable error in judgment will tend to produce a rule that covers a large number of instances (as all rules do) as if the other instances resemble the precipitating instance, even though the other instances are in reality more different from the immediate event than the rulemaker initially imagined.[16] To give just one example, when the American law of defamation was dramatically changed in 1964 in the United Stated Supreme Court case of *New York Times v Sullivan*,[17] a decision that has somewhat influenced the law in various other countries, the case before the Court was one in which the plaintiff's reputation was almost certainly not damaged at all, in which the offending publication had essentially no presence in Alabama (the state in which the suit arose), and in which the entire motivation behind the lawsuit was to punish the newspaper for its liberal views about racial integration. But the rule that emerged out of that case—the requirement that the falsity of the accusations be knowing and intentional if the plaintiff is a public official—plainly encompasses a wide range of far different circumstances, often with perverse results.

5. The Occasions of Interpretation

The previous section has focused on the occasions on which law and legal regulation initially arise, but many of the same considerations apply to legal interpretation under existing laws as well. Under one view, one most prominently associated with Ronald Dworkin, every act of legal application is an act of interpretation.[18] For Dworkin there are no easy cases in any conventional sense, and what we perceive as easy cases are in reality the product of the full interpretive enterprise.[19]

Such a view of interpretation seems counter-intuitve. For one thing, Dworkin's conclusion is parasitic on Dworkin's own Dworkinian theory of law in which there are few if any boundaries between what the positivist thinks of as legal rules and the larger array of social, political, and moral rules, principles, and norms.[20] In addition, however, Dworkin's picture denies the phenomenology by which some instances of law application appear easy, straightforward, and mechanical, while others are uncertain and troublesome. And although I acknowledge that the distinction I draw is in some sense stipulative, I believe that most of us most of the time reserve the idea of interpretation for those instances in which we perceive a quandary, or a difficulty, or a problem.[21] To say that I interpret the stop sign as ordering me to stop seems highly counter-intuitive. Maybe at some level I am indeed interpreting it, but it seems far more

[15] See Schauer and Zeckhauser 2007: 68-87.
[16] See Schauer 2006; Schauer and Zeckhauser 2011.
[17] 376 U.S. 254 (1964).
[18] Dworkin 2006, 1986, 1977.
[19] Dworkin 1986: 350-54.
[20] Dworkin 1984: 263-71.
[21] See Schauer 1992.

consistent with ordinary usage and ordinary understanding simply to say that I am understanding it, just as I understand other straightforward uses of language.

Even if we reserve the idea of interpretation for those instances in which there is a difficulty of some variety, however, it remains useful to distinguish among four types of difficulties. One, the most standard example in the law about, and the commentary on, interpretation, is the rule that is linguistically vague or ambiguous with respect to some application. Hart's "No Vehicles in the Park", when applied to bicycles, roller skates, and baby carriages, for example, is a rule of this variety.[22] Although, following Hart, the rule and its language may have a settled meaning at the core, at the fringes, or at the penumbra, we remain uncertain about the rule's application to some number of other instances. At that point many of the debates about legal interpretation come to the fore,[23] and here we encounter questions about whether we should refer to a ruler's purpose, or instead to the actual intentions of its actual drafter, or perhaps instead to the outcome that would be the best policy, all things considered. But although these are different approaches to what the law should do in cases of linguistic indeterminacy in a governing statute, regulation, or constitutional provision, the basic idea is that the occasion for interpretation arises in the first place from the phenomenon of linguistic indeterminacy.

Second, and perhaps it just a variant on the first, is the case of open texture. Open texture, the writings of too many American legal academics notwithstanding, is not a synonym for vagueness. Rather, as Friedrich Waismann, who identified the phenomenon with the German word *Porosität*, later translated by W.E. Kneale as open texture, maintained, open texture is the possibility of vagueness of a previously non-vague term when confronted with a previously unimagined application.[24] Open texture is not vagueness. It is the ineliminable possibility of vagueness of even the most non-vague term. Using an example from J.L. Austin,[25] we might understand "goldfinch" as a non-vague term, such that there are necessary and sufficient conditions for determining whether a bird was a goldfinch or not. But if we were then confronted with a creature that we believed to be a goldfinch, but which then proceeded to explode, or to quote Virginia Woolf, we simply wouldn't know what to say. That is open texture, and because old legal rules must confront the modern world, it is a common interpretive problem in law.[26]

Third is the case, and perhaps it is not interpretation at all, of factual and not legal indeterminacy. It is a requirement of the United States Constitition, for example, that the President be thirty-five years of age. But perhaps some potential President's birth certificate is unclear, or has been lost. Here we know what the rule requires, but we must "interpret" the facts in order to determine whether the rule has been followed or violated. And although such issues may rarely surface in appellate courts, they are of course the principal concern of trials,

[22] See Hart 1958.
[23] In the context of American law, see Nelson 2011.
[24] Waismann 1951: 117-44, Schauer 2013.
[25] Austin 1946: 148-87.
[26] As with, for example, the question whether patents may be obtained for laboratory created living organisms, a phenomenon plainly neither extant nor envisaged at the time that most patent laws were enacted. See *Diamond v. Chakrabarty*, 447 U.S. 303 (1980).

where interpreting conflicting facts and conflicting accounts is a large part of what trial courts are designed to do.

Finally, and perhaps most interestingly, is the case of a rule that gives a clear answer, but where for some reason the clear answer is morally, politically, or otherwise unacceptable. Dworkin's favorite example of *Riggs v. Palmer*[27] is a case of this type. The applicable rule in the New York Statute of Wills was not unclear. It was not vague, ambiguous, or in any other way linguistically indeterminate. But it was equally clear that the rule when applied according to its terms would allow Elmer Palmer, who had murdered his grandfather, the testator, in order to claim the inheritance, to obtain the inheritance. So although the law, conventionally understood, clearly dictated an outcome, it was equally clear that the outcome so dicated was morally appalling. Under Dworkin's view, avoiding the rule was still a case of application of law, but to legal positivists of a certain variety, the result was simply an example of the fact that making a decision often involves more than the law. Indeed, for Hans Kelsen[28] and Joseph Raz,[29] it always involves more than the law.

Each of these four varieties of legal quandary presents interpretive problems, but the problems they present are different, and the resources applicable to solve those problems will vary with the nature of the problem. This is not the occasion to offer an interpretive theory for each of these, but until we distinguish the four we cannot even begin to do so. More fundamentally, and more directly related to the theme of this paper, it is important to distinguish the conventional question about how we should interpret the law from the less visible but more fundamental question of when we should set forth on the enterprise of legal interpretation. These are the occasions of interpretation, and just as we need to know the occasions of law in order to understand when and how law exists, so too do we need to understand the occasions of interpretation to know when, how, and why law is interpreted.[30]

References

Austin J.L. 1946, "Other Minds", *Proceedings of the Aristotelian Society* (Supp.), XX, 148-87.

Davis W. 2014, "Implicature", in Zalta, E. (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/archives/fall2014/entries/implicature.

Dworkin R. 1977, *Taking Rights Seriously*, London: Duckworth.

Dworkin R. 1984, "A Reply by Ronald Dworkin", in Cohen, M. (ed.), *Ronald Dworkin and Contemporary Jurisprudence*, Totowa, NJ: Rowman & Allenheld, 247-300.

Dworkin R. 1986, *Law's Empire*, Cambridge, MA: Harvard University Press.

Dworkin R. 2006, *Justice in Robes*, Cambridge, MA: Harvard University Press.

[27] 22 N.E. 188 (N.Y. 1889).
[28] Kelsen 1967.
[29] Raz 1979.
[30] An earlier version of this paper was delivered to the Argentinian Association of Legal Philosophy, in Salta, Argentina, on August 22, 2017, and I am grateful for audience comments and questions on that occasion.

Ginsburg, T., Elkins, Z. and Simmons, B. 2013, "Getting to Rights: Treaty Ratification, Constittional Convergence, and Human Rights Practice", *Harvard Journal of International Law*, LIV, I, 61-95.

Grice, H.P. 1989, *Studies in the Way of Words*, Cambridge, MA: Harvard University Press.

Hart, H.L.A. 1958, "Positivism and the Separation of Law and Morals", *Harvard Law Review*, LXXI, IV, 593-629.

Hart, H.L.A 2012[3], *The Concept of* Law, 1961, Bulloch, P.A., Raz, J. and Green, L. (eds.), Oxford: Oxford University Press.

Kahneman, D., Slovic, P. and Tversky, A. 1982 (eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.

Kelsen, H. 1941, "The Law as a Specific Social Technique", *University of Chicago Law Review*, IX, I, 75-95.

Kelsen, H. 1957, *What is Justice? Justice, Law, and Politics in the Mirror of Science*, Berkeley, CA: University of California Press.

Kelsen, H. 1967, *The Pure Theory of Law*, trans. M. Knight, Berkeley: University of California Press.

Nelson, C. 2011, *Statutory Interpretation*, New York: Foundation Press.

Plous, S. 1993, *The Psychology of Judgment and Decision Making*, New York: McGraw-Hill.

Raz, J. 1979, *The Authority of Law: Essays on Law and Morality*, Oxford: Clarendon Press.

Reyes, R.M., Thompson, W.C. and Bower, G.H. 1980, "Judgmental Biases Resulting from Differing Availabilities of Arguments", *Journal of Personality and Social Psychology*, XXXIX, 2-12.

Schauer, F. 1991, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life*, Oxford: Clarendon Press.

Schauer, F. 1992, "The Occasions of Constitutional Interpretation", *Boston University Law Review*, LXXII, 729-46.

Schauer, F. 1995, "Giving Reasons", *Stanford Law Review*, XLVII, IV, 633-59.

Schauer, F. 2006, "Do Cases Make Bad Law?", *University of Chicago Law Review*, LXXIII, 883-918.

Schauer, F. 2008, "A Critical Guide to Vehicles in the Park", *New York University Law Review*, LXXXIII, 1109-1135.

Schauer, F. 2013, "On the Open Texture of Law", *Grazer Philosophische Studien*, LXXXVII, 195-213.

Schauer, F. and Zeckhauser, R. 2007, "Regulation by Generalization", *Regulation and Governance*, I, 68-87.

Schauer, F. and Zeckhauser, R. 2011, "The Trouble with Cases", in Kessler, D. (ed.), *Litigation versus Regulation*, Chicago: National Bureau of Economic Research, University of Chicago Press, 45-70.

Searle, J.R. 1969, *Speech Acts: An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.

Sport Diver, 2010 (April).

Tversky, A. and Kahneman, D. 1973, "Availability: A Heuristic for Judging Frequency and Probability", *Cognitive Psychology*, V, 207-32.

Tversky, A. and Kahneman, D. 1974, "Judgment under Uncertainty: Heuristics and Biases", *Science*, CLXXXV, 1124-1130.

Tversky, A. and Kahneman, D. 1981 , "The Framing of Decisions and the Psychology of Choice", *Science*, CCXI, 453-58.

Tversky, A. and Kahneman, D. 1986, "Rational Choice and the Framing of Decisions", *Journal of Business*, LIX, 251-78.

Waismann, F. 1951, "Verifiability", in Flew, A.G.N. (ed.), *Logic and Language: First Series*, Oxford: Basil Blackwell, 117-44.

Wyoming, 2014, "Rules of the Supreme Court of Wyoming", in *Wyoming Court Rules*, Eagan, MN: Thomson Reuters.

Constitutive Rules

John R. Searle

University of California, Berkeley

Abstract

Regulative rules regulate preexisting forms of behavior, constitutive rules make possible new forms of behavior. They constitute the phenomena they regulate. Brute facts can exist independently of any institutions. Institutional facts require pre-existing institutions, which consist of systems of constitutive rules. Constitutive rules create new forms of reality, with new powers, they typically require language, and they are the basis of human civilization.

Keywords: *Constitutive Rules, Regulative Rules, Institutional Facts, Brute Facts, Deontic Powers.*

There is an intuitive distinction between two kinds of rules: those that regulate antecedently existing behaviors and those that constitute new forms of behavior and thus regulate the very behavior that they constitute. It is natural to think of these types of rules as regulative and constitutive. I think that terminology which I first introduced in the 1950s is useful and I believe it has stuck.

I do not know who was the first philosopher to make this distinction but I believe the first published version was in an article by John Rawls in the *Philosophical Review* for 1955.[1] Rawls' article "Two Concepts of Rules" distinguished not between two types of rules but between two concepts, two types of thinking about rules. I think it is natural to construe Rawls as a forerunner of the idea that these are not just concepts about rules but that they in fact mark different types of rules. Examples of the distinction are easy to come by. So, for example, the so-called "rule of the road", according to which people in the United States drive on the right hand side of the road is a regulative rule. Why? Because the activity of driving exists independently of this rule; the rule regulates an antecedently existing activity. The rules of chess, on the other hand, do not just regulate, but they constitute the activity they regulate. So, the rule that says the King moves to any adjacent square, one square at a time, looks like a regulative rule, but in fact taken as part of the whole system it is one of the rules that in their totality constitute the

[1] Rawls 1955.

Argumenta 4,1 (2018): 51-54
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.sea

game of chess. If you do not follow these rules, or at least a sufficiently large subset of the rules, you are not playing chess.

It is furthermore natural to think of these rules as having characteristically different syntactic forms. The regulative typically takes the form of an imperative, "Drive on the right!", for example; the constitutive rules, not so obviously but I think nonetheless apparently, take the form "x counts as y" or "x counts as y in context c". So, moving one square *counts as* a legal move for the King, it would not count as a legal move for the Knight, for example.

Within any system of constitutive rules it is typical that some perform a role that is purely or almost purely regulative. Thus, the United States constitution is a system of constitutive rules and together those rules constitute the structure of the United States government, but within the system of constitutive rules there are some regulative rules, the famous First Amendment for example regulates the power of government with regard to freedom of speech and other basic human rights: "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for a redress of grievances."

The logical structure of the constitutive rules gives them enormous power. The first thing to notice is that the structure of the constitutive rule allows for its recursive iteration in such a way that you can build one rule on top of another more or less indefinitely. For example, uttering a certain sentence of English counts as making a promise, but making certain sorts of promises counts as undertaking a legal obligation, but undertaking certain sorts of obligations counts as getting married. Notice what occurred in the previous structures: on the bottom level x1 counts as y1, but y1 is equal to x2, and that counts as y2, y2 then is equal to x3, and that counts as y3 and so on up. This gives the system enormous power.

In order to understand the behavior of the constitutive rule I have to introduce another distinction that I first introduced in 1964[2] and that is the distinction between *brute facts* and *institutional facts*.[3] Some facts exist independently of any institution. The fact that the Earth is 93 Million miles from the Sun is a brute fact—it is true, we need the institution of measurement to state the fact, but the fact needs to be distinguished from the statement of the fact. The fact stated, the actual distance between Earth and Sun is a brute fact requiring no institution. The fact that I am a Professor of Philosophy at the University of California at Berkeley, however, is an institutional fact. It requires a rather elaborate institution of universities and professorships and constitutive rules of those institutional structures in order that this fact should exist. Institutional facts are always applications of the form of the constitutive rule and institutions such as money, property and government are systems of such rules. Sometimes an institutional fact can be created informally without a preexisting constitutive rule. For example, a group might treat someone as their leader and thus count her as the leader without a preexisting rule for leader selection. This presumably is one way institutions can evolve.

Constitutive rules have some further remarkable properties, which for the sake of brevity I will simply list.

[2] See Searle 1964: 50.
[3] Anscombe 1958.

1. By creating a class of institutional facts, they create a new reality. This reality is in part ontologically subjective because it only exists by way of human acceptance or agreement (Searle 1995: 69), but at the same time, statements about such facts are or can be epistemically objective.[4] So it is an institutional fact that the piece of paper in my hand is a $20 bill, but that is an epistemically objective fact, it is not just my opinion that it is a $20 bill, it is an actual objective institutional fact. The ontological subjectivity of the existence of the fact does not prevent statements about it from being epistemically objective (Searle 1995: 9-10). This is why we can have objective sciences of domains that are ontologically subjective such as Economics.

2. These new ontological domains create new powers. The facts that Donald Trump is president of the United States and that I am an American citizen involve systems of powers. Those powers are what I call *deontic* powers, that is they are rights, duties, obligations, authorizations, permissions etc. These powers are crucial in the existence of human civilization because they give us *desire-independent reasons* for acting. Having made a promise to write this paper, I have created an obligation to write it and that obligation gives me a reason for writing it which is independent of my other inclinations. Without this, we have no human civilization. We are distinct from other animals in our capacity to create these deontic powers. As far as I know, no animal has the capacity.

3. Constitutive rules require language. My dog Tarski is very intelligent and I can train him to follow various regulative rules such as to come when I call him and obey various other simple commands, but I cannot teach him to play chess or vote in elections. Why not? These phenomena only exist insofar as they are represented and the representing relation requires something more sophisticated than the simple forms of symbolism that he is capable of. This is an arguable point and I will not argue it here, but the general thesis that I wish to advance is that constitutive rules can only function insofar as they are represented linguistically. A creature cannot have a system of constitutive rules without something approaching a human or humanlike language.

4. The application of constitutive rules in institutional facts constitutes human civilization.[5] This is the chief way in which we differ from other social animals. It is customary and correct to see the human species as continuous with other species and all sorts of human forms of behavior are quite common, lots of other species have pair-bonding and other family-like forms of behavior. But no other species has the system of obligations, rights, duties and responsibilities that humans have created. No other animals have government, private property, marriage or universities. My dog, as I pointed out earlier, is very intelligent and along with other dogs he is a pack-animal and easily cooperates and moves as a member of a pack of dogs, but the pack of dogs will not get together to form a government, a university or even a primary school. Why not? In order to do these things they would have to be able to represent things in the form "x counts as y in c" common to constitutive rules, but to do that they would have to have a

[4] Searle 1995.
[5] Searle 2010.

system of representation vastly greater than their canine barks. In order to do that they would have to have constitutive rules.

Language is the fundamental human institution in that other institutions—money, property, government etc.—require language but language does not require them. For the creation of the institutional fact that I made the statement that snow is white I need only be a competent speaker of English. The semantics are sufficient to create the fact that the statement was made. But to create the institutional fact that you are president or that war is declared I need something more than semantics. We use semantics to create facts that go beyond semantics. And these facts have the structure of the constitutive rule.

References

Anscombe, G.E.M. 1958, "On Brute Facts", *Analysis*, 18, 3, 69-72.

Rawls, J. 1955, "Two Concepts of Rules", *The Philosophical Review*, 64, 3-32.

Searle, J.R. 1964, "How to Derive Ought from Is", *The Philosophical Review*, 73, 43-58.

Searle, J.R. 1969, *Speech Acts. An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.

Searle, J.R. 1995, *The Construction of Social Reality*, London: Allen Lane.

Searle, J.R. 1998, *Mind, Language and Society. Philosophy in the Real World*, New York: Basic Books.

Searle, J.R. 2010, *Making the Social World. The Structure of Human Civilization*, Oxford: Oxford University Press.

Searle on Normativity and Institutional Metaphysics

William Butchard and Robert D'Amico***

* *University of Central Florida*

** *University of Florida*

Abstract

In *Speech Acts* Searle argued for a version of philosophical naturalism by, in part, replying to G.E. Moore's famous claim that naturalism, if it included any evaluative claims, would be clearly fallacious. We make the case that Searle's reply was not the disaster it is sometimes claimed to have been. In our discussion we pay special attention to Searle's introduction of such key concepts as brute facts, institutional facts, and constitutive rules. We also make a broader case for the 'constitutive' connections as central to Searle's often misunderstood metaphysical views. We intend to show at least that Searle has an account of normativity that, while in a sense constructivist, is both naturalist and realist.

Keywords: Normativity, Realism, Naturalism, Social ontology, Constructivism, Grounding.

1. Introduction

In the last two decades Searle's work on social philosophy, or what could also be called his metaphysics of institutions, has been widely discussed and critically assessed.[1] As seems characteristic of Searle's views on various topics, those debates have concerned whether Searle's overall position coheres. Specifically, it is common to hear that Searle's constructivism about institutional phenomena is inconsistent with his philosophical naturalism.[2]

Though we have also entered into these debates (Butchard and D'Amico 2011, 2015), we now intend to reconsider Searle's earlier philosophical career where the groundwork for these views was first laid. In fact, we intend to

[1] See Corcoran 2001, Ruben 2015, Hindriks 2009, 2013. See also Searle 2004 for responses to critical assessments of *The Construction of Social Reality*.

[2] For example, Ruben 2015 characterizes Searle as a "constructivist" because the institutional domain is in effect built out of mental states and physical states. We reply to Ruben in our concluding remarks.

Argumenta 4,1 (2018): 55-68
ISSN 2465-2334

© 2018 University of Sassari

DOI 10.14275/2465-2334/20187.but

reexamine in this essay Searle's first major philosophical work *Speech Acts: An Essay in the Philosophy of Language* (Searle 1969).

In that work Searle defends his version of philosophical naturalism by replying to G.E. Moore's argument in *Principia Ethica* (Moore 1993, originally published in 1903) that naturalism with respect to evaluative terms can be shown to be fallacious. In his reply, Searle introduces the key concepts of brute facts, institutional facts, and constitutive rules. Though the discussion of Moore is not based on a close reading, and Searle only sketches these important concepts, we are confident that these passages hold a key to understanding Searle's mature views.

Our aim is two-fold. We make a case for Searle's defense of naturalism with regard to normativity (as properly understood). We also argue in defense of his concept of constitutive rules, which we see as critical to his entire metaphysics of institutions.[3] More specifically, we aim to show that Searle defends an account of normativity that is both realist and naturalist.

We discuss his appeal to constitutive rules as a special case of his appeal to constitutive connections generally. Searle takes the term "constitutive" seriously enough (from the point of view of metaphysics) for his picture to ensure both that the phenomena of which real things are constitutive of are themselves real, and that if everything constitutive of some phenomenon is natural, then that phenomenon is itself natural. There is more about Searle's appeal to constitutive connections in the final section of this article.

What is key to Searle's account of institutional obligations is the fact that the attitudes constitutive of them are normative in character. Depending on how one understands Searle's account there could be a threat of circularity. We argue that the sense in which normative attitudes are constitutive of actual obligations, on Searle's account, does not involve circularity. Nor do we think it commits him to a form of mind-dependence that could be taken as a threat to realism. Thus, if the attitudes are both real and natural, the actual obligations of which they are constitutive are real and natural as well.

2. Moore on the Naturalistic Fallacy

In *Speech Acts*, Searle embraces the view that normative properties are real and thus part of philosophical naturalism. As we will explain, given the function of constitutive rules, Searle can be said to be realist about social phenomena and specifically institutions (though of course these phenomena are mind-dependent, or what he also calls "representation dependent"). These considerations also ensure that he is a realist about obligations in institutional settings.

Before launching into this discussion, it may be helpful to clarify some terminology. For some of this clarification we are just stipulating matters, while with respect to other points we set aside controversies that cannot be addressed in this article. However, our intent is to emphasize that Searle's embrace of "naturalism" should not be confused with what is currently called "physicalism," for instance; though physicalism is a specific version of naturalism as we understand it.

[3] Searle's metaphysics in philosophy of mind is a version of "anomalous monism" since the natural world is not law-like throughout and there cannot be a science of the mental. We cannot address this large topic here, but in addition to the articles of ours cited above, for a brief discussion of the differences between Searle and Davidson on anomalous monism see D'Amico 1997.

Toward clarifying what conception of naturalism Searle is defending, we believe it to be a broad commitment to understanding the world by locating phenomena within the framework supplied by the sciences, again with the term "science" being understood broadly. Scientific naturalism does not commit one to any particular metaphysics; rather, it allows for different metaphysical positions, e.g., physicalism versus dualism.

To say either that mental properties are wholly grounded in physical properties or that they are merely caused by them is to locate mental properties (explanatorily speaking) within the scientific worldview. Scientific naturalism is thus compatible with an appeal to *a priori* principles that complement empirical data. Such principles need not be limited to mathematics; they can include analyses that address conceptual connections generally. Hence, physicalism, including versions of it that rest crucially on *a priori* considerations, counts as naturalism.[4] For purposes of grasping Searle's project we also stress that naturalism does not concern itself with how various phenomena are known in the first place; rather it is concerned with how various phenomena can be worked into the scientific worldview.[5]

It is worth noting in this context that Searle considers himself neither a physicalist (strictly speaking) nor a dualist. While he is not a dualist in philosophy of mind, Searle does not think mental properties supervene on physical properties with the strict necessity required by physicalism. But he does not think that such a position commits him to property dualism. His reason is that he thinks property dualism presupposes a substantive pre-theoretical mental/physical dichotomy, and for Searle there is no such dichotomy. In contrast, as regards our topic, he *does hold* that normative properties supervene on mental and physical properties with strict necessity.[6] Thus if his theory of mind merits the label "biological naturalism" then his account of normativity should be classified as naturalistic.

Given this conception of naturalism, there arises a famous problem that has been the focus of much of Searle's philosophical career. On the face of things, a core element of daily thought and language concerns the normative aspects of states of affairs. But it has been argued that the existence of normative features of the world is not—and could not be—part of the scientific worldview. Hence, we have the following problem: Either normative features do not belong to the natural world or there are no normative features.

Thus while at first sight normative statements are factual and thus part and parcel of the natural world, taken broadly, the study of the natural world in the sciences has been thought not to allow for normative or evaluative facts. Of course, this problem should be stated carefully. There are obviously legal systems and there are facts such as those concerned with rules. These facts would certainly be part of the natural world. But it is possible to consider these cases as not

[4] We agree with much of how David Copp (2003, 2015) characterizes naturalism epistemically and the dispute about normativity. Also Copp outlines what he calls a social theory of ethics as a response to how ethical properties could be understood naturalistically. Whether the positions of Searle and Copp are on the same track is simply too much to cover in this paper.

[5] See Searle 1992: 1, 85.

[6] Searle of course does not use the language of supervenience in *Speech Acts* and he does not make serious use of it in his later writings. In fact, he explicitly rejects the term as both misleading and unnecessary (Searle 1992). But, as we will argue, his appeal to constitutive connections secures strong supervenience.

involving *normative facts* as opposed to just *normative attitudes*. These cases may well be wholly accounted for in terms of psychological or perhaps even physical facts and thereby ultimately not require anything more than normative attitudes.[7] Thus we could put the challenge in this manner: Were the sciences to provide a vast catalogue of objects and properties that make up the world, that catalogue would not include normative or evaluative facts or properties (though it likely would include normative attitudes as psychological states).

This line of thought seems to leave broadly two options. First, there is the option of retaining the naturalist outlook and adopting an anti-realist position with regard to normativity. Second, one could commit to the existence of normativity but treat it as a non-natural property.

We cannot discuss details of these options here, but we want to emphasize before we begin the exposition that Searle rejects both options. In other words, Searle aims to be a realist about normativity. On his view, normative properties exist, and normative properties are natural properties. Understanding Searle in this fashion then makes clearer why Moore's claim that such a combination of views constituted a fallacy had to be answered.

If Searle's view is that moral properties are real and natural, then his account must in some way explain moral properties and perhaps in turn explain why they may seem mysterious. This task will require Searle to be clear about what is expected of him in defending both realism and naturalism. But in understanding Searle's position it is important to stress that by natural property Searle does not mean a property within some existing science. His only requirement is that it be explainable in terms of properties in the sciences. He thus does not commit to the view that there could be a science of normativity (in parallel fashion Searle argues there cannot be a science of the mental).

An example of Searle's position on naturalism can be found in the way he treats the category of furniture.[8] Samples of furniture are physical objects, and whatever properties those objects have are also natural properties. But of course there is no science of furniture, and it would be foolish to think there ought to be such a science. Furniture will not be found in the catalogue of objects and properties as prepared by the sciences. But, Searle insists, it would be wrong to conclude that therefore one should be an anti-realist about furniture. Given the established sciences, including such fields as physics, psychology, and economics,

[7] In explaining Searle's naturalism we say that he aimed to defend the reality of normative *facts* as part of the world. When we appeal to the notion of a normative *attitude* we simply wish to point to an even more modest position that we think is also compatible with the scientific outlook. For instance, an emotivist such as A.J. Ayer could well embrace that there are normative attitudes as a fact of human psychology while also maintaining that there are no moral properties. We want to thank the anonymous reviewer and David Copp for pointing out the need for further clarification on these matters. We address our use of the term "constitutive," albeit briefly, in our discussion of David-Hillel Ruben and specifically in footnotes 15 and 16.

[8] "Some philosophers point out that we don't have a science of behavior for the same reasons that we don't have a science of furniture. We couldn't have such a science because that aren't any physical features that chairs, tables, desks, and all other items of furniture have in common that would allow them to fall under a common set of laws of furniture. And besides we don't really need such a science because anything we want to explain—for example, why are wooden tables solid and why does iron lawn furniture rust?—can already be explained by existing sciences" (Searle 1984: 73).

furniture is neither mysterious nor non-natural. It is ultimately a higher-level feature that can be accounted for in terms of familiar mental and physical properties, all of which are natural.

In turning to Searle's discussion of Moore in *Speech Acts*, the first matter we encounter is the way in which Moore put his demonstration of the naturalistic fallacy. Moore claimed he had demonstrated that one could not analyze ethical statements in terms of naturalistic statements.[9] He held that derivation of an evaluative statement from a factual base requires that the term "good" be synonymous with some descriptive term *F*, making it analytic that if something is *F*, it is good.

Moore's requirement on derivation, in brief, is that there be analytic bridge principles connecting evaluative statements with descriptive statements. Moore then argued against the possibility of such a derivation by way of what he called an "open question argument." If the term "good" and *F* were to bear an analytic connection, then when presented with something that is *F,* it could not seriously be said to be open to question whether it is also "good." Moore of course had in mind *a priori* analytic connections analogous to those between the words "bachelor" and "unmarried male."

But Moore's open question argument runs into problems. It depends on a sense of "openness" about our response to such sentences. It also depends on a sharp distinction between normative and non-normative statements that Moore struggled to provide.

Contemporary philosophers are apt to respond to this argument by citing the work of Saul Kripke (1980) and Hilary Putnam (1973) on natural kinds. Kripke and Putnam have given us reason to think that Moore's argument fails because the meaning of natural kind terms is not exhausted by *a priori* analytic definitions, making it possible for two expressions to be synonymous even though their synonymy is not transparent to ordinary (i.e., empirically uninformed) speakers. And there is the related point that just as water's being identical to H_20 is what Kripke calls an *a posteriori* necessity—making water's H_20 nature an open question to ordinary folk—it may be that normative kinds are identical with natural kinds even though the identity is accessible only *a posteriori*. Hence Moore's central demonstration of the fallacy collapses.

Of course, Searle's discussion of Moore has none of the fancy objections about natural kinds we just discussed. In fact, Searle is brief about what is wrong with Moore's demonstration. He offers examples in which he thinks the derivation goes through, but his discussion, as we will see, engages later thinkers more directly than it engages Moore. We believe that Searle's point about Moore concerns what constitutes the factual basis of the derivation. Before discussing this issue, however, we want to look at some of Searle's examples.

3. Searle's Argument for Descriptive-Evaluative-Entailment

Searle begins his response to the Moorean tradition with examples from J.O. Urmson. Searle wants to explain why such examples have been thought to be impossible or to raise problems for naturalism. Urmson discusses, for example, fruit grading terms introduced by the British Ministry of Agriculture and Fisheries. Take the statements:

[9] Moore concedes in his "A Reply to My Critics" (Moore 1942: 535-677) that his efforts to clearly make a distinction between normative and naturalistic claims were a failure.

- This apple is Extra Fancy Grade.
- This apple is *A*, *B*, and *C*.

Urmson holds that the first claim is not entailed by the second because it is evaluative while the second is descriptive. Searle replies that the ministry saw itself as *defining* "Extra Fancy Grade." Given their definition, Searle points out, it is analytic that any apple that is *A*, *B*, and *C* is Extra Fancy Grade. Thus the statement "The apple is *A*, *B*, and *C*" entails "The apple is Extra Fancy Grade" (Searle 1969: 135-36).

Searle notes that it is fundamental to later Moorean arguments to assert a gulf between the *meaning* of an evaluative term and the *criteria for its application*. But, he says, once you identify "Extra Fancy Grade" apples as ones that are by definition "*A*, *B*, and *C*," there is no such gulf (Searle 1969: 134). He grants that the *characteristic illocutionary force* of the utterance of "The apple is Extra Fancy Grade" is *to grade something* while the characteristic illocutionary force of the utterance of "The apple is *A*, *B*, and *C*" is *to describe something*. But he argues that this difference is not sufficient to show that the proposition expressed by the utterance of the one sentence fails to entail the proposition expressed by the utterance of the other. We have to distinguish, he says, the characteristic illocutionary force of an utterance from its propositional content, as well as the closely related pair *use* and *meaning*. (Searle 1969: 136).

To help make Searle's more effective strategy clearer, it is worth looking briefly at an anti-realist variant on Moore's argument. The work of R.M. Hare best represents a classic anti-realist position in agreement with Moore's claim of a naturalistic fallacy.[10]

Hare showed how anti-realists can agree with Moore about there being a naturalistic fallacy without accepting Moore's specific argument against naturalism. Hare's own version of the open question argument is as follows. If naturalism were true, then evaluative terms would be descriptive terms. But their being descriptive would not make sense of the fact that we often *commend* objects *for being a certain way*. If "good" were a property term that attributes the feature that forms the basis of an evaluative judgment, then our statements of the form "x is good because it is *F*"—which commends something for having feature *F*—would be read is "x is *F* because it is *F*," which does not commend x for being *F*. Hence, descriptivism (which is implied by naturalism) mistakenly reduces statements of the form "x is good because it is *F*" to statements of entirely different form, which would make it impossible to say things of the sort we normally say with evaluative terms. Hare's argument thus shows that Moore's so-called open question turns in the end on a gap between the meaning of an evaluative term and its descriptive criteria.

We have spelled out this version of the naturalistic fallacy because it rests on the distinction between meaning and criteria. While Searle uses Urmson's examples, his diagnosis as to why the derivation has appeared impossible was meant to apply to Hare as well. The diagnosis, once again, is that philosophers in this camp fail to distinguish *use* from *meaning*, leading to the mistaken distinction between the meaning of an evaluative term and its descriptive criteria. He thus

[10] Hare (1952) mounts his defense of Moore's naturalistic fallacy argument in Chapter 5 of *The Language of Morals*. Hare (1972) also responds to Searle concerning the relationship between meaning and use, a topic we have to set aside.

addresses Urmson and Hare more directly in this section than he does Moore (Searle 1969: 134-35).

Searle does not discuss the is-ought question in this part of *Speech Acts*. And we think it is at best unclear that his examples, as he describes them, show that it is possible to derive ought-statements from descriptive statements. Is it analytic that one ought to accept arguments whose premises entail their conclusion? Is it analytic that one ought to seek out apples that are *A*, *B*, and *C*? If the answer to these questions were clearly yes, then Searle's examples might be said to show that it is possible to derive an ought-statement from a set of descriptive statements. But these claims are not clearly analytic. One might doubt, for example, that when the British Ministry of Agriculture and Fisheries stipulates that "Extra Fancy Grade" applies only to apples that are *A*, *B*, and *C* that "Extra Fancy Grade" retains its ordinary normative meaning. If they are not analytic, Searle has simply not made a case for any principle that would make the derivation possible. Thus, as far as Searle's examples go, there may well remain an is-ought gap, even given his theoretical framework.

However, as we will now discuss, Searle identifies in a later part of *Speech Acts* constitutive connections between the attitudes people have and genuine obligations they have as a result within institutional settings. We think these connections more plausibly bridge the is-ought gap.

4. Searle's Is-Ought Argument

Once again, when Searle turns to the task of actually producing a deductive argument to answer Moore, Searle appeals to analytic bridge principles. In this case, the principles make possible a derivation of statements containing the term "ought" from a descriptive expression containing the term "promise." But unlike the examples noted above, Searle now can appeal to the following plausible analytic principle: If one promises to X then (baring extenuating circumstances) one is under an obligation to X. Searle does not rely on the principle's intuitive appeal. Rather, he sketches a theory of institutional reality that treats promising as an institution and explains why promising places someone under an obligation. He then takes it as analytic that if one is under an obligation to X, then one ought to X.

Searle does not try to provide a reductive analysis of normativity to something wholly descriptive in the sense that no normative terms are required to specify it. He thus makes free use of normative terms in identifying attitudes that contribute to the factual base for the derivation. On Searle's view, obligations can arise by virtue of our collective acceptance of certain actions as placing one under an obligation. For example, we accept that to utter "I promise" is to issue a promise and hence to place oneself under an obligation (it being analytic, as noted, that to issue a promise is to place oneself under an obligation). And since it is analytic that one ought to do what one is under an obligation to do, a person should do what he or she has promised.

Crucial to Searle's view is that the collective acceptance consists in the "we-intentions" of individuals. Each of us has an intention with the content "We intend that to utter 'I promise' is to make a promise." No individual's we-intention is enough for an utterance to amount to promise. But when enough people (and perhaps key people) have the intention, the acceptance counts as collective and the utterance really is a promise and really does place the person uttering it under

an obligation, in which case we have (what we call) a normative fact rather than a mere collection of normative attitudes.[11]

In brief, even if the concept of a promise is normative, and our individual we-intentions have normative content, it takes multiple individuals with such attitudes for genuine obligations to arise. Thus we can include normative terms in the descriptive base as long as they do not introduce the normative facts that arise from the individual states described in that base. Searle, we believe, thus relies implicitly on a bridge principle in which there are occurrences of normative terms—"obligation" and "promise"—that do not introduce normative facts, but are followed by an occurrence of "obligation" that does introduce a normative fact. For example:

- If a community collectively decides that a certain action brings one under an *obligation* by virtue of constituting a *promise*, then the action really does bring one under an *obligation*.

Searle's aim is to identify the institutional conditions under which someone is placed under an obligation, but the conditions he identifies are not supposed to explicate the very idea of an obligation (or of any particular kind of obligation). He identifies some of the conditions under which obligations may be realized in the institutional world, but the conditions are properly identified by means of normative terms because they include individual attitudes with normative content. In brief, Searle shows how institutional facts can give rise to obligations; he does not try to show that the concept of an obligation (or of a particular kind of obligation) is analyzable in terms of strictly descriptive concepts. We will say more to clarify Searle's explanatory project below in our response to Ruben.

The form of the bridge principle suggests an answer to the worry that Searle's use of normative terms in specifying the derivative base makes that base not wholly descriptive thereby making his derivation circular. Searle's derivation base is wholly descriptive in one sense and not wholly descriptive in another. It is wholly descriptive in the sense that it does not introduce *normative facts*. Rather, it introduces the normative attitudes of individuals, which, as we have stressed, are not themselves normative facts but may ground them. The descriptive base is not wholly descriptive only in the sense that it properly includes normative terms. But as long as those terms do not introduce normative facts, Searle will have derived normative facts from a set of statements that do not themselves state normative facts.

Notice that the bridge principle above explains how one might *know* that a person ought to do X. If one knows that a person uttered, "I promise to do X" and knows that we have collectively decided to treat the utterance of this expression as a promise, then, one knows that the person ought to do what was promised. The reason, as Searle argues, is that it is analytic that promising places one

[11] When Searle wrote *Speech Acts* he appealed broadly to the notion of the community in his account of language use and he also used expressions that Searle would later develop at greater length as a whole doctrine of social phenomena and "we-intentions." This later position (Searle 1990) turns out to be complex and we must set it aside here. However, we tried to show (Butchard and D'Amico 2015) that Searle's account of we-intentions cannot be reduced to internal, individual causal states that some recent authors have called "incentives." The reader can also consult Margaret Gilbert (2007) for her challenges to Searle's we-intentions as against Gilbert's similar notion of what she calls "plural subjects."

under an obligation and that one ought to do what one is under an obligation to do. In this way one is in a position to know that the person ought to do X.

Moore's argument that the normative is not reducible to natural properties (as shown by the open-question argument) cannot take hold against Searle. Searle does not explicate the normative "ought" in terms of something non-normative. But Searle does show how to derive a statement of normative fact from other sorts of statement. Searle's derivation simply does not require the kind of reductive analysis Moore associated with naturalism.

5. Searle on Constitutive Rules

Searle's aim in the philosophical section of *Speech Acts* is grander than just replying to Moorean arguments against naturalism. He wants to show that his appeal to institutional facts establishes a broad conception of how normativity is part and parcel of the natural world and thereby part and parcel of philosophical naturalism. He states that it is at least sometimes a matter of institutional fact (not brute fact) that persons have commitments, obligations, duties, or responsibilities.

The point we are stressing now is that Searle's dispute with Moore fundamentally concerns the question of whether the concept of naturalism ought to be restricted to brute facts alone. Searle thinks that naturalism cannot be so restricted, but he does think that institutional facts are layered upon brute facts. However, to fully explain the relevant idea of layering requires access to bridge principles in the form of constitutive rules.

As Searle (1969: 184) put it, "No set of brute facts can entail an institutional fact without the addition of at least one constitutive rule." Searle rejects here a problematic restriction on which facts can figure in the deductive base in the derivation of ought claims. But, as we noted, in Searle's discussion of Urmson and Hare, Searle does not directly address whether *moral* ought-statements can be derived from descriptive statements. Though Searle's discussion leaves it unclear whether the way he derives ought-statements in institutional settings will have a counterpart in non-institutional settings, we think he has, at least, shown that the possibility of such a wider derivation should not be ruled out.

Searle clearly notes that he has not addressed the moral question directly.[12] But our aim throughout has not been to show that Searle has a completely worked out account of ought-statements. Rather our point is that his treatment of such statements in institutional settings does, plausibly, show that it is possible to derive ought-statements from descriptive statements and hence to fit some such statements into an ambitious naturalist picture of the world.

For our purposes, Searle's discussion in *Speech Acts* is the first appearance in his work of the distinction between brute facts and institutional facts, and then the further device of constitutive rules.[13] Searle rejects a picture of the world

[12] In *Speech Acts* (1969: 187), Searle takes himself to have shown, by way of example, the falsity of the general thesis that an "ought" claim cannot be derived from an "is" claim, and as having undermined the linguistic thesis upon which the more specific moral thesis has rested. He thus thinks that, in light of his institutional examples, a special argument is needed for the narrower thesis that a moral "ought" cannot be derived from an "is".

[13] These notions are of course not original with Searle. In *Speech Acts* (1969) Searle credits the distinction between brute and institutional facts to Elizabeth Anscombe and in "What is a Speech Act?" (1965) Searle recognizes John Rawls as the source for the distinction between regulative and constitutive rules.

containing only brute facts capable of being known through experience, and language tracking only these facts. This picture, according to Searle, makes it very difficult to understand ethics and esthetics and has led to the implausible view that such statements are mere psychological reports, contrary to their initial appearance. "It cannot be said that the implausibility of these ways of dealing with the problems posed by ethics and esthetics has been any bar to their popularity, but their popularity is at least evidence of the power of this picture" (Searle 1969: 50-51). He then adds:

> Leaving aside the question of the status of statements in ethics and aesthetics, which are controversial areas anyway, there are many kinds of facts, and facts which obviously are objective facts and not matters of option or sentiment or emotional at all, which are hard, if not impossible, to assimilate to this picture (Searle 1969: 51).

Searle has in mind institutional facts, as opposed to brute facts. He first marks the distinction entirely by way of examples. Examples of brute facts are gravitational force, location in three-dimensional space, or physiological processes such as digestion. Institutional facts are exemplified by marriage ceremonies, legislative actions, economic exchanges, or legal trials. We believe Searle's point of contact with Moore is that he thinks the open-question argument appears persuasive only if we restrict our attention to brute facts.

After giving the examples of brute and institutional facts, Searle draws our attention to the important additional notion of constitutive rules:

> These "institutions" are systems of constitutive rules. Every institutional fact is underlain by a (system of) rule(s) of the form "X counts as Y in context C." Our hypothesis that speaking a language is performing acts according to constitutive rules involves us in the hypothesis that the fact that a man performed a certain speech act, e.g., made a promise, is an institutional fact. We are not, therefore, attempting to give an analysis of such facts in terms of brute facts (Searle 1969: 51-52).

We first want to point out that there is nothing in the distinction between constitutive and regulative rules to exclude a naturalistic account of normativity. To say that institutions, or institutional facts, exist only because we adopt certain rules does not make them unreal, but of course it does make them mind dependent. If it is a fact that the bank is closed, then there is a system of material objects that exists independently of our minds, and there are individuals (all of who exist independently of one another's minds) who are in collective agreement that some objects in the world have a certain status. The latter condition requires the existence of minds, but it does not make the statement about the condition of the bank any less factual.

But there is another matter we want to address. At the beginning we stressed that Searle's position strikes many as best understood as an anti-realist view. As we noted earlier, Ruben argues that Searle's constructivism leads to anti-realism about institutions. Ruben makes crucial appeal in his article to the notion of supervenience. He argues that Searle's account of social reality is best understand as the specification of the details in the supervenience of the social on individual persons, their mental states, and physical objects. But for such an account to be successful, Ruben holds, it must *explain why* the supervenience holds. He does not think Searle's constructivism supplies such an explanation. Ruben points out,

correctly we think, that supervenience is not itself a *determinative* connection, so it is not the sort of relation that could *explain* institutions in terms of persons and material objects. Supervenience, he says, always obtains because of some further connection, and Ruben holds that Searle does not identify that further connection.

For instance, Ruben comments on the following statement by Searle: "The central span on the bridge from physics to society […] and the decisive movement on that bridge […] is the collective imposition of function" (Searle 1995a: 41). Ruben argues that if this passage is intended as an account of the relation between levels that thereby explains the supervenience, it has failed. Ruben reaches that conclusion because the existence of collective imposition must belong to one of the levels, whether one thinks of it as introducing the summation of individual attitudes or as something at the social level. Ruben's criticism is that collective imposition therefore cannot be a bridge across levels. Ruben says:

> The main point is only this: constitutive rules and the collective assignment of function (whether identical or not) must be either members of the set of I or set S or set M, for they must be placed somewhere and these exhaust the options of where to place them (it is obvious that neither are at the brute physical level) (Ruben 2015: 137).

Ruben's point, again, is that Searle's appeal to the collective imposition of function as the grounds for the existence of institutions must place it at one of the levels—I (the set of institutional facts), S (the set of social but non-institutional facts), or M (the set of mental facts)—but then it does not serve as a bridge. According to Ruben, Searle's picture merely directs our attention to a co-variation between levels without explaining that co-variation, and hence does not introduce an explanatory bridge principle as realism would require. Ruben's conclusion then is that Searle's *construction* of institutional reality amounts, unacceptably, to brute supervenience, and hence leaves the door open for anti-realism. We agree with Ruben that brute supervenience (of any strength) is a threat to realism.

We want to stress that if this is a problem for Searle's account of institutional reality, then it is a problem for his naturalism about normativity. The reason is simple: Naturalism about normativity entails that the normative supervenes on the natural. But if supervenience is not an explanatory notion, and yet is the key relation between the normative and the non-normative that defines Searle's naturalism, then his naturalism is likewise subject to Ruben's criticism.[14]

We believe that Ruben identifies a very important issue in this discussion. But we think that it is not particular to Searle's metaphysics. Ruben has in effect raised a broad concern about philosophy as a mode of inquiry. When philosophers engage in analysis they are not, as one might think they are, looking for strictly modal connections such as supervenience. They are more often looking for explanatory connections. Searle (1992: 125) himself distinguishes "causal

[14] Hare (1989: 66-81) claims that the correct use of moral terms supervenes on the use of descriptive terms, but he does not say that it supervenes with the strict necessity of the supervenience entailed by naturalism, of the sort that says that moral properties strongly supervene on non-moral properties. In other words, his account of moral terms involves weak supervenience, which says only that any descriptive duplicates within the same possible world are evaluative duplicates. Hare thus does not commit himself to the stronger thesis Moore associates with naturalism, namely the one that requires analytic bridge principles that the open question argument then rules out.

supervenience" from "constitutive supervenience". The terms "causal" and "constitutive" are both explanatory terms. When Searle says that some rules are constitutive of an activity rather than imposed on a prior activity, he is pointing to a non-causal explanatory role they play.[15]

We think it reasonable to take philosophy as in the business of seeking out constitutive connections. We also think it reasonable to understand such constitutive connections in terms of non-causal explanatory notions such as metaphysical grounding or even real essence (though we seriously doubt Searle would go that far). Nevertheless, Searle's use of the term "constitutive" in these contexts suggests an answer to Ruben's proper question of why the supervenience relation holds: I supervenes on M (or S) because M (or S) is constitutive of I.

Moreover, constitutive terms are used in cases in which the phenomena, and hence the levels, are in an important sense not wholly distinct: the higher levels (levels being sets of property instantiations) exist *wholly in virtue of the existence of the lower levels*. Thus the account we are discussing has it that institutions are not wholly distinct from the material objects and collective acceptances that ground them. But then to return to Ruben's criticism, the claim that the levels are non-distinct in this manner makes it unproblematic that collective acceptance has to be located at one of the levels. It is, on Searle's account, located at a level that fully ensures the existence of the derivative institutional level. Grounding, therefore, provides a bridge between levels thereby explaining why the higher levels strongly supervene upon lower levels. Ruben's criticism that the location of collective acceptance at one of the levels excludes it as a bridge between levels appears to rest on the assumption that the levels are wholly independent and hence not *inherently bridged*. But the view that there are constitutive connections between levels is precisely the view that the levels are non-independent.

We stress by way of clarification that the term "constitutive" can be used in two ways. It can be used to introduce explanatory conditions that are *essential* to some phenomenon, and it can be used to introduce explanatory conditions that are not essential to but nevertheless fully account for (i.e., *metaphysically ground*) some phenomenon.[16] Searle appears to use the term in both ways, and seeing both uses is key to understanding his naturalism. He appears to use the term in its essentialist sense to give an account of institutional reality. Institutions, on his account, essentially involve collective intentionality. But in his account of

[15] We use "constitutive" as a generic term for a class of explanatory notions that came into focus in the last half-century as philosophers began to suspect that strict identity was not always adequate for capturing connections across domains (e.g., identity rules out the so-called multiple-realizability of the mental or the coincidence of the statue and the hunk of clay that constitutes it). This family of concepts includes, but is not limited to, realization, constitution, essence, and recently grounding. The term "constitutive," as we use it, indicates that these notions are non-causal. Hence, Searle's makes a distinction between causal supervenience and constitutive supervenience. We also understand the concept as introducing notions that are not strictly modal. For a discussion of the inadequacy of strictly modal terms for capturing such notions (specifically that of essence), see Fine 1994.

[16] Being red, e.g., is not *essential* to having a color, but it may, in some particular case, fully *ground* it. By contrast, having a surface feature of a certain reflective sort is essential to having a color. The contrast in this example is not supposed to imply that essentialist connections exclude grounding connections. Rather, we think it is reasonable to hold that having a surface feature of a certain sort grounds having a color and that this is closely related to the essentialist connection.

normativity—his derivation of an "ought" claim from an "is" claim in institutional settings—he appears to use the term in the second sense. That is, he appeals to constitutive connections between attitudes and obligations in institutional settings, but he does not claim that obligations (or any particular kind of obligation) are essentially tied to those attitudes. In other words, he does not argue that the concept of obligation at work in his institutional metaphysics is different from any concept of obligation that may apply outside of institutional settings. In brief, it is his view that facts about collective intentions can ground real obligations.

Thus we take Searle's use of constitutive connections as introducing explanatory connections in metaphysics, as opposed to treating explanation as restricted to the work of the sciences. But if this sort of effort at explanation is in some way improper or inadequate on Ruben's view, then we think that such a broad criticism has the unfortunate consequence of driving philosophy wholly into deflationary or eliminativist projects, excepting, perhaps, cases in which they appeal to identity. Of course, this kind of explanation that we defend and think is exemplified in Searle's long effort to rescue a rich sense of naturalism is both somewhat messy and even tentative. But that will be the case in any non-deflationary account. We are under no illusion that there are not serious and weighty problems that need to be addressed, among them the status of non-causal explanations. But it would be a loss to abruptly dismiss Searle's effort at securing and defending a robust naturalism. If we have made our case, then there is a version of naturalism compatible with the sciences and common sense, but which also preserves the critical autonomy of philosophy.

Bibliography

Butchard, W. and D'Amico, R. 2011, "'Counting As' a Bridge Principle: Against Searle Against Social-Scientific Laws", *Philosophy of the Social Sciences*, 41, 455-69.

Butchard, W. and D'Amico, R. 2015, "Alone Together: Why 'Incentivization' Fails as an account of Institutional Facts", *Philosophy of the Social Sciences*, 45, 315-30.

Copp, D. 2003, "Why Naturalism", *Ethical Theory and Moral Practice*, 6, 179-200.

Copp, D. 2015, "Explaining Normativity", *Proceedings and Addresses of the American Philosophy Association*, 89, 48-73.

Corcoran, K. 2001, "The Trouble with Searle's Biological Naturalism", *Erkenntnis*, 55, 307-24.

D'Amico, R. 1997, "Impossible Laws", *Philosophy of the Social Sciences*, 27, 309-27.

Fine, K. 1996, "Essence and Modality", *Philosophical Perspectives*, 8, 1-16.

Gilbert, M. 2007, "Searle and Collective Intentions", in Tsohatzidis, S. (ed.), *Intentional Acts and Intentional Facts*, München: Springer, 31-48.

Hare, R.M. 1952, *The Language of Morals*, Oxford: Oxford University Press.

Hare, R.M. 1970, "Meaning and Speech Acts", *The Philosophical Review*, 79, 3-24.

Hare, R.M. 1989, *Essays in Ethical Theory*, New York: Oxford University Press.

Hindriks, F. 2009, "Constitutive Rules, Language and Ontology", *Erkenntnis*, 71, 253-75.

Hindriks, F. 2013, "Collective Acceptance and the Is-Ought Question", *Ethical Theory and Moral Practice*, 16, 465-80.

Kripke, S. 1980, *Naming and Necessity*, Cambridge, MA: Harvard University Press.

Moore, G.E. 1942, *The Philosophy of G.E. Moore*, Schilpp, P.A. (ed.), Evanston: Northwestern University.

Moore, G.E. 1993, *Principia Ethica*, Baldwin, T. (ed.), Cambridge: Cambridge University Press (originally published in 1903).

Putnam, H. 1973, "Meaning and Reference", *Journal of Philosophy*, 70, 699-711.

Ruben, D.H. 2015, "Beyond Supervenience and Construction", *Journal of Social Ontology*, 1, 124-41.

Searle, J.R. 1965, "What is a Speech Act?", in Black, M. (ed.), *Philosophy in America*, London: Allen & Unwin, 221-39.

Searle, J.R. 1969, *Speech Acts: An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.

Searle, J.R. 1984, *Minds, Brains, and Science,* Cambridge, MA: Harvard University Press.

Searle, J.R. 1990, "Collective Intentions and Actions", in Cohen, P.R., Pollock, M. and Morgan, J., *Intentions in Communication*, Cambridge, MA: MIT Press, 401-15.

Searle, J.R. 1992, *The Rediscovery of the Mind*, Cambridge, MA: MIT Press.

Searle, J.R. 1995a, *The Construction of Social Reality*, New York: The Free Press.

Searle, J.R. 1995b, "Reply to Critics", *Philosophy and Phenomenological Research*, 57, 427-58.

Searle, J.R. 2004, "Realism Reconstructed: A Reply", *The Philosophical Forum*, 55, 275-80.

Searle, J.R. 2010, *Making the Social World*, New York: Oxford University Press.

Meta-meta-institutional Concepts? A Tale on Schwyzer and the Force of Technical Ends (Live from Ruritania)

Guglielmo Feis

University of Milan

Abstract

The paper is a critical analysis of Hubert Schwyzer's idea of meta-institutional concepts. First, I isolate a presupposition in Schwyzer's example of chess as ritual. I then show how Schwyzer's idea of meta-institutional concepts is far from being the endgame in the research on levels of institutionality. In fact, we can iterate on meta-institutional concepts. Schwyzer's idea has to face an infinite regress. I try to avoid such a regress by introducing the concept of technical end of game. A game defines its own terminal status. People playing the game can then attach different meanings to this norm-dependent terminal status. Hence, meta-institutional concepts are not conditions of possibility: they belong to pragmatics, not to an extra ontological layer. I conclude the paper sketching a classification of games based on different kinds of technical ends and I advocate the need to adopt a pluralistic conception on games and practices.

Keywords: Meta-institutional Concepts, Technical Ends, Constitutive Rules, Social Ontology, Hubert Schwyzer, Infinite Regress.

The paper is a critical analysis of Hubert Schwyzer's idea of meta-institutional concepts.

The chief example of a meta-institutional concept is that of *winning in a game*. Games are different and have different rules. Nonetheless, they share the concept of victory that is present in most rulebooks. As Roversi (2014) says, we use the same word to say "I *won* at chess" and "I *won* at bridge", but it will be (pragmatically) odd for us to reply "Yeah, I won, too, yesterday at bridge" to a friend that says "Yesterday I won at chess".[1]

[1] Schwyzer (1969) first focused on victory. Miller (1981) talked about meta-institutional concepts discussing promises. The most recent research wholly dedicated to meta-

Argumenta 4,1 (2018): 69-83
ISSN 2465-2334

© 2018 University of Sassari

DOI 10.14275/2465-2334/20187.fei

There seem to be two alternative options to account for the "new dimension" of meta-institutional concepts:

1) We *add a layer* to our ontology of the institutional domain: the obvious option is to dedicate a *new level* to meta-institutional concepts;
2) We *do not add a layer* to our ontology of the institutional domain and try to account for this "new dimension" relying on other tools.[2]

The goal of this paper is to criticize the standard view of meta-institutional concepts, denying that we need to follow option (1). Against this route I shall offer both methodological criticisms (§§ 1-2) and a conceptual problem based on the idea of properly defending a new foundation for the institutional domain (§ 3).

I criticize the view according to which meta-institutional concepts are conditions of possibility of the institutional realm (Lorini claims that straightforwardly;[3] Roversi's approach seems closer to the one I will develop in the paper)[4]. I then build a positive proposal highlighting how the meta-institutional theory failed in distinguishing between the *end of a practice* and the *meaning we assign to the end of a practice* (§§ 4-5).

The thought driving the *pars destruens* is that when we claim there is an extra level or a background or some presupposition we also need to prove that that is the ultimate level. The leading question is whether the discovery of the alleged meta-institutional concept is the last step in our research of the background and the presuppositions of constitutive rules.

The thought driving the *pars construens* is that we need not look back or beyond or *meta constitutive rules* to discover their background but rather *inside* them. They will tell us all we need to know as far as how the game will end.

1. Every Saga Needs a Prelude (and some Methodological Remarks)

The starting point of the paper emerges as soon as John Searle's theory of social reality interacts with Schwyzer's (1969) paper. Both Roversi (2014: 202, 214) and Lorini (2014: 127) say that the Searlean picture of institutional reality needs to be developed and enriched. Lorini thinks we need a new level and favors the approach involving meta-institutional concepts (option 1 from the section above). I will say, on the other hand, that we only have to add an *element* (not a

institutional concepts I know of is that of Lorini (2012, 2014) and that of Roversi (2007, 2014, 2015).

[2] Giuseppe Lorini developed the first option claiming that meta-institutional concepts are conditions of possibility of the institutional, see Lorini (2012, 2014). Corrado Roversi focused on the second option and proposes a pragmatic analysis of the speech acts we use to describe the supposed meta-institutional dimension.

[3] "In this paper I extend Searle's theory of institutional facts arguing that a further level is needed for the investigation of the structure of institutional reality: *the level of meta-institutional concepts*. The *meta-institutional concepts are concepts that go beyond (Greek: metá) the institutions of which they are conditions of possibility*" (Lorini 2014: 127, second *italics* are mine; see also 129).

[4] Roversi's idea of *parainstitutionality* comes close to what I am arguing for in this paper. To paraphrase him (2014: 211) "Technical ends thus figure as the fundamental layers on top of which parainstitutional usages rest".

whole new *level*) in the ontology, viz. the concept of technical end, i.e. the terminal point of a practice. Now to work.

First, the analysis involving meta-institutional concepts applies to a wide variety of contexts, such as games (chess, tennis, football, videogames, hide and seek, etc.), social practices (e.g., rituals, social exchanges such as greetings, etc.), institutions (promising, marriage, governments, elections, money, etc.) and many more. I propose to generalize and talk about rule-constituted element(s) (RCE for short) rather than rules, practices, games. 'Element(s)' shall be neutral enough to cover all these different grounds. Nothing hinges on this terminological choice. If you prefer 'entities' as more neutral than 'element', no problem with that. '*Constituta*' might be another option.[5] My talking of RCE is just a reminder that we need to focus more on the pluralities involved in our discussions of constitutive rules and not only rely on chess and other "cheesy and easy" well-known examples.

RCEs (in the plural) owe their existence to a set of rules and they are wholly dependent on the set of rules that generates them. They are the products (which can be entities belonging to different ontological categories) of sets of constitutive rules that determine them. There are two important features that are brought to the fore by reflecting on this use in the plural.

First, we say that constitutive rules are constitutive "as a whole". In order for a constitutive rule to produce its effect(s) we need to embed it into a context, into a "totality of the rules" (Lorini, Roversi, Conte and Searle all agree on that). Some of these rules might be constitutive as well, some might be regulative rules or meta-rules (higher-order norms).

Because of this embedding and because of the fact that other rules refer to rules-constituted elements (e.g., in the case of technical legal gaps) we have different relationships and interactions between rules belonging to the set of constitutive rules. That's why (constitutive) rules are constitutive "as a whole" or "in context".

If so, then speaking of the products of constitutive rules favoring the singular over the plural would be misleading. It suggests that the thing in question is the product of a single rule. There would be no difference between the lone element and "the whole" and the latter would be the former under a different name. For example, let's imagine that we have this single rule:

(CR 1): "paper banknotes count as money if they are issued according to procedure P";

Does it create money as a singular rule constituted-element? Obviously not. In order to create money by means of constitutive rules, we would need many other of rules, such as:

(RR 2): "you can't deny a payment that is made by means of valid forms of money".

Such a (RR 2) rule allows you to pay a coffee with a credit card—assuming it counts as valid money, something that might require us to add a third rule in the system—or even with a fifty dollars banknote.

On the other hand, money (and most of other entities created by means of constitutive rules) is not an amorphous mass. It is complex and articulated into

[5] Żełaniec (2013, 39) introduced the term *constituta* to refer to something constituted by constitutive rules or one such rule. According to that usage money is a *constitutum* (in the singular) and so is volleyball.

components (such as single coins, banknotes, cheques, credit cards and so on) which, down to a certain level, are constituted, too, by means of other rules of the same system to which both (CR 1) and (RR 2) belong.

This is why I prefer to say that money (and the like) is (a system of) *RCEs* (highlighting the plural). This will help us, or so I hope, never to forget that money (and the like) is a structured and articulated whole, a whole whose elements are constituted by some from among the rules that define money as a whole.

The same goes with chess. Chess (singular) is a (system of) RCEs. All the constitutive rules that define the different pieces interact. If a pawn reaches the rank furthest from its starting position, the player who managed to do that ought to promote such a piece turning it into a different. The promotion rule (art 3.7.5.1-3.7.5.3. of the FIDE laws of chess) determines the chess RCE of a pawn and plays a similar role to (RR 2) above, helping us constructing a toy-model for chess as a (system of) RCEs.

There is an extra reason to talk about RCEs (in the plural) that is due to the fact that the plural reminds us of a variety of phenomena. Chess, money, marriages, football games, institutional entities, etc. are all RCEs in as much as they are the product(s) of set(s) of norms that include at least one constitutive rule.

Sometimes, theories of constitutive rules seem to be designed to account for one and only one phenomenon: chess. Such theories are those that preach saving the phenomena but consider only the single case that makes the theory work: chess. However, it could be fruitful and instructive to adopt a wider selection of examples and include such games as tennis, blackjack, *Magic: the Gathering*, go, scrabble (*stems* anyone?)[6] or even computer games. Philosophy of computer games is arguably a subdiscipline of philosophy and game design a discipline on its own. None of these (sub)disciplines are mentioned in the chess-dominated debate on meta-institutional concepts.[7]

The sets of rules most often specify circumstances in which RCEs are over. In chess, it is a specific situation: checkmating the opponent's king (or a stalemate). Other forms of technical ends may include time length, points, physical distances or a combination of these elements (see §4 for a sketch of a classification). This does not rule out that certain RCEs can go on indefinitely. Tennis admits such a possibility. The same was true for soccer when draws in in-or-out matches (say the final phase of a World Cup) caused the game the be repeated. The case of some RCEs that might not have a crystal-clear technical end (e.g., non-competitive playing, training) are briefly considered below (§5). Think about a computer running the set of rules of some RCEs.[8] Rather than 'winning'

[6] In competitive Scrabble you are interested in "bingo stems". You make a bingo and get extra points when you use all the letters on your rack. Six letters stems are combinations that easily make a bingo if you add other letters. An example is the "TISANE" stem that combines with all letters but J, Q and Y.

[7] See, e.g., Myers (2003), Moore (2011) and Sageng, Fossheim & Larsen (2012). The criticism that a certain theory of constitutive rules has been developed on chess only has been voiced already. See Leonardi (1983: ch. 3—for a discussion of soccer as constituted practice) or Feis (2014—on different game phenomena such as doping and simulation—and Feis 2018).

[8] People working in AI are keen on considering and modelling a wider array of phenomena than philosophers. We have championship in which algorithms play chess and Computer Olympiad which features a wider set of games. Another option to represent

or 'losing' I will talk about "reaching the technical end of the game". Once we disclose this, different possibilities of interpreting this wholly rule-dependent terminal point open up: it could be a victory in a competitive game, it could be a ritual, it could be [add your favorite]. Maybe this is too a "static" or "computational" approach in analyzing the phenomena. At the present stage I can re-emphasize that, at least, though "static", "computational", "formalistic" or the like, at least it allows to catch phenomena in the plural.

This has a further consequence for the issue we are debating. If we agree on RCEs-driven pluralism, then we have to evaluate at least as dubious and incomplete all proposals that defend meta-institutional concepts based on chess only.[9] A chess-based theory of meta-institutional concepts is a hastened generalization.

In what follows I will try to analyze Schwyzer's famous case of chess as a game and as a ritual. In so doing, we will try to write down the rules of both chess-as-a-game and chess-as-a-ritual. They are going to be two systems of RCEs and we shall try to highlight some of their features with toy-models similar to that presented above for money as (CR 1) and (RR 2). The *desideratum* would be: being able to write down a list of the (constitutive) rules constituting the game in order to implement it on some sort of simulation or computer program.[10]

Another advantage of proceeding as if we were coding (i.e. building software) the practice is that if we are able to implement at least a toy model simulation of different systems of RCEs we will gain a common ground on which to evaluate different games and practices.[11]

Here, I shall be assuming that we are dealing with practices and activities that *do have* an end. The objection that I omit never-ending practices does not apply to my argument as the advocates of meta-institutional concepts consider chess and victory most of the time. Further, alleged never-ending practices have an end: Shéhérazade is not immortal (so her supposed counterexample of a never-ending narration reaches and end). Bastian saves *Fantasia* screaming the name of his mother (should he fail in doing so the other possible outcome is that *Fantasia* is destroyed). If you think of miming games or mimesis-as-make-believe games, there is still some teleological end.[12]

the running of RCEs could be Ethereum's smart contracts. See De Filippi & Wright (2018) for a primer.

[9] Even when other sports are mentioned, I am not aware of authors who tried to unveil the meta-institutionality of other sports or games performing an operation similar to that of Schwyzer (1969) with chess.

[10] Both a chess videogame (enter your favourite) or a sort of NetLogo simulation will be ok. NetLogo is a programming language that creates agent-based models and simulation. A quick overview can be found here: https://en.wikipedia.org/wiki/NetLogo. In NetLogo we will be forced to distinguish two kinds of agent that are relevant for Schwyzer's arguments: priests and non-priests.

[11] Models will be toy-models, sure. But chess as the only test for meta-institutional concepts (in general) is a toy-model as well.

[12] Żełaniec (2013), who flirts with the idea of meta-institutional concepts, is probably the author that stressed the most the importance of teleology and human nature/biology as far as constitutive rules are concerned, see esp. chs. 7-8. See also the importance of teleology in Roversi (2014)'s aim-oriented meta-institutional concepts. This teleological end and the concept of acting-in-function (what Amedeo G. Conte calls *nomotropismo*) of a rule can be used to deal with the supposed counterexamples. I partially did that as far as training is concerned in Feis (2014), see also §5 of this paper.

Further, I think that methodologically it is a mistake to rely exclusively on ordinary language argumentation when discussing institutional concepts. One would have to explicitly decide such questions as: Which language from among the many on Earth? Which usage? Based on which *corpus*? Yet this is seldom done. No matter how naïve your institutional consideration is going to be, in a couple of sentences we can reach a point in which language usage does not conform to the intended theory you try to argue for by way of ordinary language. Besides, do institutional documents such as written laws, rules in a manual or even speeches by legal authorities or leaders count as ordinary language (think about property, marriage or even elections)?[13] If we assume that normative language, in particular the legal discourse, is some sort of semi-planned language, ordinary language philosophizing on that should be regarded as at least suspicious. Something similar will follow if we have an artefact theory of law: it is the (ordinary) language of the artefact maker and of the community that makes the law, but there is no guarantee that that is the ordinary language.[14]

2. The Exclusivity Rule or the Dirty Little Presupposition in the Construction of Meta-institutional Concepts

Meta-institutional concepts date back to Schwyzer's (1969) thought experiment meant to show that practices have no built-in instructions concerning how we should enact them. In his (in)famous Ruritania case, Schwyzer tries to show that we can have two kinds of chess practice constituted by the *same rules*, call them the FIDE-chess[15] and RITE-chess: the first practice is a competitive game, the second a religious ritual.

There is a lot of room for arguing that RITE-chess has different rules from FIDE-chess (*contra* Schwyzer's assumptions—oddly enough, all the philosophizing on Schwyzer's paper never focused on that). Neither Schwyzer nor the ad-

[13] The mistake here is that of reasoning on what is taken as an institutional level (that of meta-institutional concepts, constitutive rules and RCEs) running ordinary language arguments. Some hold the view that institutional languages such as legal language are not so close to ordinary language. If there's a point in claiming that an institutional language is to a degree different from everyday language, then an institutional language has its own usage and "grammar". I think those applying ordinary language or linguistic considerations to institutional matters need to deal with this point first. I actually share Haas' (1953: 514) worries on ordinary language helping "creating a philosophical muddle, though a muddle expressed in correct English". Funnily enough, legal philosophers insisting on distinguishing institutional language as peculiar adopt ordinary language philosophy narrowly conceived (in practice: analytic philosophy done by informal arguments based on those linguistic usages which are apt to prove your point).

[14] In a Legal Positivist setting it is even easier to see that there is not much room for the ordinary language at all, it is the language of the authority making the law that counts (though the authority has reasons to stick to what is thought to be the standard usage).

[15] I find it at least controversial to say that FIDE rules are ordinary language, at least if we have to give some credit to the work in legal pragmatics and the idea that law is something in between an artificial and a natural language. Lorini (2014) occasionally employs ordinary language arguments based on the FIDE law, e.g., section 2.3. of the paper saying that the article 10.1 of FIDE presupposes winning. See also Lorini (2014: 134). Such claims need testing in a Computer Olympiad and other scenarios in which non-human players execute sets of rules (see fn. 8).

vocates of meta-institutional concepts ever carried out an explicit—"extensional" if you will—characterization of the rules.

Let's present some reasons why the rules are *not* the same in RITE-chess and FIDE-chess. If we carefully consider Schwyzer's example we find that there is something different from FIDE-chess in his ritualized version (I do not know if this counts as a phenomenology of the game, Schwyzer offered us no rule-book, just his intuitions). Think about reconstructing a model of FIDE-chess and RITE-chess in some sort of regulated and coded environment (e.g., NetLogo). The world of RITE-chess features two different kinds of players (priests and ordinary folks, i.e. non-priests) whereas in FIDE-chess all players are equal.

We can clearly see this difference if we refer to the "Blasphemer!" part of Schwyzer's story. In Ruritania ordinary people, who are not entitled to play chess, are sanctioned (in the paper the exclamation "Blasphemer!" expresses this sanction). In RITE-chess, in fact, there are *two kinds of agents*. If a non-priest tries to enact the chess rules a sanction occurs. In FIDE-chess there is no such thing.

In order to account for this element, we need to carefully *add a rule* in our model of RITE-chess, call it the *exclusivity rule* according to which only priests may enact the chess[16] system of RCEs. If we were to program this, we would have a domain of agents (priests and non-priests), then the program would randomly pick an agent from the domain and try to have the agent execute the system of RCEs. If the agent is not a priest nothing will happen. There is no outcome of the chess-like practice, because it cannot start. If the agent is a priest, then the game starts and we will have a final stage of this game and an outcome.

According to that representation, the rules composing RITE-chess are different from those of FIDE-chess, in fact RITE-chess includes the exclusivity rule above (and it seems it does not include stalemate). In order to have Schwyzer's story going and to have a chess-like game in which there's no winning and losing but only a ritual, it is essential to have a division between priests and non-priests. It seems to me hard to deny that the exclusivity rule is part of the game. FIDE-chess and RITE-chess are not the same game because they have different rules. Nonetheless, RITE-chess and FIDE-chess largely overlap, and RITE-chess contains at least some elements of the technical end of FIDE-chess (checkmate is still there, stalemate—we do not know).[17]

Advocates of meta-institutional concepts may try to remove the exclusivity rule from the game and implement it at another level, e.g., in the social context or something else. Society in Ruritania is such that when a non-priest is playing chess he or she gets sanctioned.[18]

[16] Chess-like, if you prefer.

[17] To my knowledge, such an issue has never been investigated. For our present purposes; (i) if stalemate is not part of RITE-chess, RITE-chess is a different game compared to FIDE-chess. That's an easy win against Schwyzer as the assertion "we are using the *same* set of rules in different ways (defined by means of different meta-institutional concepts)" is false; (ii) as far as my reconstruction in terms of technical end goes, we can adopt some sort of charitable reading and restrict our attention to the aspects of FIDE-chess and RITE-chess that overlap without some Ruritanian exegesis. In this setting, both FIDE-chess and RITE-chess share the technical end according to which "the king can't be safely moved".

[18] I think that spelling out these details would bring in a whole new level of complexity that requires us to review the whole example. Even by doing this, the issue of stalemate will still be there allowing us to say the rule sets are different.

The above reconstruction allows us to see that, at the very best, Schwyzer's case is ambiguous, depending on how we realize the exclusivity rule. Advocates of meta-institutional concepts should clarify which reconstruction they are considering: where do we place the exclusivity rule? I am not aware of any discussion of that. The following table sums up the reconstruction:[19]

	FIDE-Chess	**RITE-Chess**
Agents playing the game	2	2
Type of Agents	1 (players)	2 (priests, non-priests)
Layers of the Game (context(s))	1 (technical end)	2 (technical end + social interpretation)
Sanctions	1? (violation of fair play)	2 (violation of fair play + "Blasphemer!" sanction)

As we have said already FIDE-chess and RITE-chess share a terminal status of the rule-governed activity, i.e. checkmate (stalemate needs further evaluation and will be left out of this discussion). You need a checkmate to have a winner, you also need a checkmate to evaluate the ending of a ritual. It is not difficult to expand the Ruritania phenomenology and image different conditions for RITE-chess.

You can either have RITE-chess 1, in which the ritual is in favor of Ruritania's people if Black reaches the checkmate or RITE-chess 2, in which the ritual is in favor of Ruritania's people if White reaches the checkmate. More generally: winning and losing can be associated differently with the same kind of technical end: in card game like Hearts the more points you score the worse for you, in Rummy the more you have the better.

3. A Regress Is Haunting Institutional Reality

The idea of iterating on a concept is far from new. In the analysis of constitutive rules, we welcomed meta-institutional concepts as a way to go beyond our main topic of investigation and investigate it at another level. This was our new hope.

Nonetheless, every time we add another level of inquiry, the question of whether: (i) there is only the starting level and the meta-level (i.e. the newly discovered level) or (ii) there are more levels (the starting one, the meta-level and more levels that are implied or maybe presupposed by the new one).

In particular, we need to be sure that levels are not added *ad infinitum*, otherwise our newly discovered "promised land" level might compromise the research. We do not want our new hope to become a new dope.

Advocates of meta-institutional concepts say that the true picture of the institutional domain is two-fold. I would challenge them by asking whether they have a closure rule to ensure that their foundation is the ultimate one.

I start by asking whether we have an infinite number of meta-levels of meta-institutional concepts, such that we could distinguish meta-meta-institutional concepts (M^2-concepts for short), M^3-concepts (i.e., meta-meta-meta-institutional concepts), etc., up to M^ω (and possibly beyond that). To my knowledge, in the

[19] I have included sanctions in FIDE-chess as far as fair play is concerned. This might be controversial but nothing in the argument depends on that.

contemporary research on meta-institutional concepts these questions have not (yet) been raised.

I think it is enough that, for a given theory, we can *conceive* of possible iterations of the meta-levels to hold that a theory claiming there's a single meta-level that is the foundation of institutional reality does not work (for some examples see later §5). Further, as Lorini proposes meta-institutional concepts as *conditions of possibility* of the institutional domain, we can demand that a foundation at a higher level should legitimize its status of being the ultimate level. Searle gives us reasons (that we may or may not accept) to think that all the levels we have are those he is talking about; Schwyzer and those who follow him have not provided any reasons to think that their levels are ultimate.

Lorini challenged Searle by pointing out (with Schwyzer's help) that what Searle thought to be the conditions of possibility of social reality rested on other conditions of possibility (Lorini's meta-institutional concepts). The issue is whether we can be sure that meta-institutional concepts are the last step, the ultimate ground, or, rather, we have to infinitely keep adding further levels. If adding meta-institutional concepts is going to add also meta-meta-levels and meta-meta-meta-levels, then introducing the first meta-level will at best be useless, at worst: dangerous. I think the burden of proof that there are just two levels is incumbent on the advocates of meta-institutional concepts.

What is happening in this paper is nothing but an echoing of an argument of the Neo-Schwyzerians: they chide Searle saying "well, you forgot to mention the level that gives a foundation to the institutional reality you construct". Here I am saying nothing but "well, are you sure your newly discovered level is the ultimate level?".

4. Technical Ends: or How to Fix Meta-institutionality without Potentially Infinite Levels

Now that the main questions addressed to the advocates of meta-institutional concepts have been raised, it is time to see if we can do any better as far as the iteration problem is concerned: are there further backgrounds for technical ends? Are there meta-technical ends? What about meta-meta-technical ends?

I argue that technical ends are first level entities due to their own (as RCEs) specific constitutive rules. If my claim is correct, we will probably have to review the place of meta-institutional concepts. Meta-institutional concepts are in fact, it will turn out, meanings we assign to the technical ends that determine what we can do with our RCEs. Consequently, they are no conditions of possibility, *pace* Lorini. We need to already have our RCEs before we can ask questions on what we can do with them.

In both FIDE-chess and RITE-chess there is a common subset of rules that determines checkmate as a technical end common to both kinds of practice. Once we reach this rule-dependent status in an instance of RITE-chess or FIDE-chess, both are over: i.e. both systems of RCEs are over when checkmate happens (we do not know if RITE-chess may end with a draw, as in FIDE-chess).

We can now enter another level and discuss how to evaluate it or "use" it. Is it White winning and Black losing? Is it the other way around? Is that something new like something good forthcoming for Ruritania's citizens because of a ritual? Is it winning your life while playing against Death himself as in the film *The Seventh Seal* by Ingmar Bergman? No matter which one you choose, there is

still some core-common rule set running in the background. The meanings we assign to the technical ends (meta-institutional concepts) shall tell us how to use those different RCEs that terminate when an instance of the technical end is reached, selecting one way from among different possibilities, without being a condition of its possibility.[20]

4.1 A Technical Ends-Based Recognition

There are different kinds of technical ends of various RCEs. Technical ends define the end of an instance of the respective (system of) RCEs in that, once a certain status, defined by a (sub)set of rules on which the RCEs depend, obtains, the instance of the given RCE is over. We have already seen this for checkmate (and stalemate). Card games may end when there are no more cards you are allowed to draw (e.g., solitaire).

Other technical ends may include a reference not only to a status of the activity but also to specific events: a bell's ringing (as in fighting), the fact that a certain amount of time has passed (for example in a football match), someone doing something for the first time inside the constituted space (think about being the first to cross the finish line in a race). Chess itself has an alternative technical end as well: checkmate or stalemate (though we most often reason without the latter).

Yet other technical ends may include spatial considerations in defining the relevant event or status. Think again about the case of running a marathon. Most racing in the Olympic games is distinguished by length (100 m, 400 m, and the like), i.e. a spatial condition. Of course, the length determines the event (crossing the line and finishing the race) as well as a status (being the first in the race).

You may add and combine these elements.[21] Think about Formula One racing which has an alternative technical end: (i) it works as a normal race—a marathon with cars—but it has to last *less than two hours*; (ii) if the race lasts more than two hours, the two hours project a sort of virtual finish line and the race is over: the positions reflect those which players occupied at the two hours mark. Points might be another form of a technical end (*Pictionary* or the funny way in which in tennis points are aggregated in games and sets).[22]

[20] Torrengo (2016) and his institutional externalism might connect to my proposal, though his focus is not on the background of constitutive rules but on institutions in general. Another remark that could be regarded as somewhat technical-end friendly is that of Azzoni (1988: 44): "the ontology of *lusus* is completely determined by the rules of the *ludus*. There can be no mediation of an extraludic reality because, as far as the *ludus* is concerned, there is no extraludic reality".

[21] Considerations above (§4.1) are guidelines for such a taxonomy. Paradigmatic properties may be whether the technical end is highly impacted by material features, say weight, strength or height; whether the technical end features spatial elements or only rule-related elements (e.g., points); etc. Some issues pertaining to the materiality of constitutive rules have been explored in Feis (2018). Sandbox, open-end games, and training in view of the best performance are the major challenges for that. Sandboxes and open-end games nonetheless have different events of gameplay that feature some technical ends, see Karhulahti (2013).

[22] Quick comparison with volleyball. No, the fifth set is not a tie-break. You can have tie-breaks in all sets, not only in the fifth. In all big tournaments (with the notable exception of the US-open) you can't have a tie-break on the fifth set.

The game ends when the technical end defined by its rules has been reached. At this point it is all up to us to decide what to do and how to interpret it. Checkmate can either be me winning and you losing, or a ritual. There is no need to specify which of the two options (or another option) we have chosen in order to instantiate a certain (system of) RCEs in so far as the RCEs share a subset of their constitutive rules that define the same technical end.

As I tried to show discussing RITE-chess, FIDE-chess and Bergman-chess: the sets of rules offer us a terminal status that is both common and prior to any further extra characterization. If we consider the case of White checkmated Black we have at least: (1) White prevailed against Black in a certain competitive game; (2) a certain ritual is completed in a positive way for people of Ruritania; (3) you saved your life. We can also invert the outcomes (1)-(3): if White checkmated Black. Call this inverted-chess (again, feel free to designate this with a different label). In order to obtain them you have to modify a rule in the core set of rules saying that if you checkmate your opponent you lose the game. Plausibly, to make inverted-chess a playable and enjoyable game you will have to add some rules that "force" each player to attack the opponents' pieces and "move towards" checkmate.

Actually, considering inverted-chess shows us that the common-core of different uses of chess presupposes only the fact that checkmate (and stalemate) are designed in such a way as to terminate the system of RCEs.

You may want to add some sort of (Aristo)telic conception to this, saying that systems of RCEs tend to reaching their technical end. I do not want to make such a commitment. We can further elaborate interesting taxonomies of technical ends, we can even dare say that all systems of RCEs have a technical end.

5. Summing up: Criticism to the Proposal and Replies

It is time to see what kind of criticism may be raised against various parts of this paper.

Let's consider the issue of meta-institutional concepts and their iteration. Meta-meta-institutional concepts seem not so absurd and, relying on the type-token distinction, we can get a handle on what is going on there. Be it as it may, let's try to offer different examples of institution-related concepts that can enter into a three-levels structure (a level, a meta-level and a meta-meta-level) instead of a two-levels one.

First, consider the case of victory (e.g., in a game). Is there a way to embed this prototypical example of a meta-institutional concept into a further level? Maybe culture might be such a meta-level for victory. Certain cultures are fonder of individualism and competition than others. Empirical game-theory investigated that and offers a base for that point. Researches on how money is split in the Dictator game exhibit different patterns of sharing across different cultures. In the Dictator game the dictator chooses first how to split a certain sum of money with the other player. The other player can do nothing but accept the dictator's choice. Rational choice theory says the dictator should take all the money for herself. In practice—and against rational choice theory—people playing the game share money most of the time (in different percentages).

Further, victory could not be the last word and the most important point of a culture. There is no need to make cross-cultural comparisons this time. Think about any debate on who's the best player in a certain sport (e.g., the Messi *vs.*

Ronaldo struggle at almost every FIFA award). Is victory all that matters? Maybe not. Style matters (that's the argument for "McEnroe was better than Borg" in tennis). Maybe being merciful and forgiving prompts your honor and social prestige more than being a cruel winner.

A second institutional-related example comes from the view of legal philosophy which Riccardo Guastini entertains. For Guastini (2012: 57) jurisprudence is the meta-level of (positive) law. Jurisprudence takes positive law as its object of inquiry and develops its doctrines starting from that. The most distinguishing claim of Guastini's conception of philosophy of law is that, given that philosophy of law has to be linguistic analysis, the objects that philosophy of law can investigate are nothing but the different doctrines of jurisprudence. Philosophy of law is, then, nothing but meta-jurisprudence (that comes in two varieties, descriptive and prescriptive, but that does not interest us here). Hence, if we reunite the levels and make use of transitivity we shall be able to see that philosophy of law is the meta-meta-level of (positive) law.

Now let's move to technical ends. Given the arguments against Schwyzer (§2-3) and the alternative proposal advocated here (§4) let's see what complaints and replies could be voiced against this analysis.

Roversi discusses an "internal, technical concept of victory" and proposes a challenge based on ordinary language we have already seen in the beginning of this paper:

> Player A says, 'Yesterday I played chess and won.' Player B replies, 'I, too, won yesterday, playing bridge.' This conversation is perfectly meaningful, but if different games had their own concept of victory, then player B's use of the '*too*' would be paradoxical, as if B had said, "Did you really take the bishop yesterday? I did the same. I, *too*, took the bishop at the airport." In the first exchange, the two players are talking about the same thing, even though this same thing (victory) is subject to different rules in one case and the other. By contrast, in the second exchange, they are referring to two entirely different situations, even though the expression they use is the same (Roversi 2014: 206, fn. 5).

The first response is that B's statement is lacking an adversative element, i.e. "I, too, won; *but*, playing bridge". I think such an adversative conjunction is there pragmatically, even if not written down in Roversi's example.[23] Admitting there is such an adversative element (i.e. *but*), the dialogue shows that we have *two different technical ends* that are peculiar, the two participants are aware of that fact and relate to each other because, functionally speaking, reaching a technical end is the last step of an instance of a system of RCEs. That is what they are using to realize that they have something in common: "You achieved the final stage in X, I achieved the final stage, but in Y".

Let's try to offer some further elements to support that line of reply. Let's see other two instances of "speeches"—i.e. following Roversi's argumentative style—in which two people refer to different instances of the same X and use an adversative. This should prove that there's no meta-institutional X. They are just two different instances of the same stuff.

[23] At least, it is there in my understanding of that example as an ordinary language conversation.

Think of "I met Tuomas, yesterday". "Oh, I, too, met Tuomas, (but) today" where A and B are referring to different Tuomas(es) and know that. There is no meta-institutional Tuomas, there are just two different guys named "Tuomas".

Think about "Yesterday I played the notes C, D, E, F, G, A, B over a C major chord" and "Oh, I, too, played the notes C, D, E, F, G, A, B (but) over an F major chord". Again, here there is no "meta-institutional C, D, E, F, G, A, B".[24]

Maybe the fans of meta-institutional concepts conceive these concepts as types that get instantiated by different tokens. We have a common type or *Ur-type* of victory that gets instantiated in many different ways in different systems of RCEs. That would be similar to cases in which in the word 'vittoria' presents two tokens (instances) of the type *t* and two tokens (instances) of the type *i*. The literature on types (sub-types) and tokens (and their occurrences) abounds with problems and difficulties that, of course, are going to affect our talking of types and tokens with reference to meta-institutional concepts.[25]

Supporters of meta-institutional concepts may reason as follows. There is a (meta-institutional) type of 'victory', whatever this might be, that gets instantiated by different sub-types, say 'FIDE-victory', 'RITE-victory' and even 'Bergman-victory'. There are some problems here we need to explore.

First, when you play two different games of a certain (sub)type, say, FIDE-chess (singular), you have multiple occurrences of FIDE-victories (plural). In a FIDE-chess tournament there are different matches that enact the same set of rules, FIDE-rules. These FIDE-rules are the same for all matches. You may say that FIDE-rules define a type and that the different games are different tokens enacting those rules. This seems also the line of reasoning of Lorini when he says that Fisher's victory, Karpov's victory and Deep Blue's victory seem to be instances of the same type. For Lorini that should be an instance of the meta-institutional type of victory.

We already have a type (FIDE checkmate technical ends turned into FIDE-victory: checkmate is winning in a competitive game) and tokens (different winnings in actual individual games). How can we accommodate in this picture the fact that FIDE checkmate and RITE checkmate are different sub-types of the same (meta-institutional) type?

Let's look at this issue backwards, i.e. we start from Karpov's token of the type "victory": we then have the FIDE-chess-victory type (i.e., checkmate as competitive victory) and meta-institutional victory as a higher order type. Here again arises the issue of victory as meta-meta-institutional concept which we have seen above discussing the regress problem (§ 3). Be it as it may, it seems that the theory of meta-institutional concepts needs some extra work before using the type-token distinction that, rather than solving all issues in one fell swoop, is likely to pose extra threats.

[24] The two actually played two different RCEs: C Ionian (*aka* the C major) musical scale and F Lydian. Here we have two scales that are constituted by different rules. The rules determine what the intervallic structure of the scale is. In the case of the major scale you have the following intervallic pattern (T stands for tone, S for semitone; a tone equals two semitone): T T S T T T S, the scale grades are the following: 1, 2, 3, 4, 5, 6, 7. Lydian would be T T T S T T S. The corresponding grades are 1, 2, 3, 4#, 5, 6, 7. C Lydian would thus be C D E F# G A B.

[25] See at least Wetzel (2008). Those debates are not often mentioned when the concepts of type and token are used in discussions of constitutive rules.

Second, even assuming that meta-institutional concepts such as victory are types that are broken up into subtypes and then instantiated differently depending on the game (FIDE-victory, type 1; RITE-victory type 2; and whatever else we may come up with) then not all the types in a given ontology are going to be meta-institutional. We cannot generalize the claim that, given that—by assumption—the type of 'victory' will be meta-institutional, all the types that receive multiple tokenized instantiations are going to be meta-institutional as well. Think about the type of ∅ (i.e. the empty set) for an example. In what way does it have anything meta-institutional about it?

Hence, we need to have extra details on why only certain concepts can be conceived of as types that admit different sub-types that are differently structured (as RITE-victory and FIDE-victory).

Third, the idea of meta-institutionality seems to fit in well with cases in which we have sorts of practice and games and, more generally, things that are executed or performed. If all types that receive multiple instantiations are going to be meta-institutional, we can challenge that view saying that there are also "conceptual types", i.e. types that are hardly executed or performed because they do not give rise to any practice. You could also use different definitions of 'meta-institutional concepts' (say Lorini's, Schwyzer's and Roversi's) as an example of different tokenizations of the same type of concepts. Note that, again, you may want to sort out this issue by saying that the conceptualization are types of a different higher order supertypes.

Fourth, as it always the case when types and tokens are mentioned, we have to consider the issue of whether there can be tokens without a corresponding type. Despite the fact that arbitrary mereological sums are most often discussed in the debates on universalism and unrestricted composition in mereology, they come in helpful in the analysis of the institutional domain. If you consider a token of the type 'university' and take a closer look at it, it's likely to be a scattered object: different departments located in different positions, students and professors that are even more scattered. Further, a university is probably composed of different kinds of objects: physical, institutional, etc. Different universities offer different degrees; some have a Ph.D.-course, some don't. Some university may have patents registered and be also a special kind of economic agent. Some include research centers. The more we get into a single token phenomenon, the harder it gets to find a type fine-grained enough to say that our beloved concrete instance of university X we are analyzing is a token of that type.

There are further challenges that we can raise to the technical end view. Training and solitaire are two possible challenges for the concept of technical end: where's the technical end of a card game solitaire? What about training (e.g., practicing your backhand in front of a wall?).

Luckily, the idea of acting-in-function of a rule (*nomotropism*) can help us in managing the cases of training. Training is parasitic on technical ends. We practice serve and volley because tennis' technical end is based on points and serve and volley (used to) help making points. We practice game endings in chess because of its technical end (checkmate and stalemate, sometimes our best option is to achieve a draw instead of a losing). And so on.

Puzzles and solitaire (and some computer games) are not a problem for a technical end-based analysis. They have a technical end defined by the correct completion of the puzzle or, in some card games, the game ends when the number of cards you can draw and manipulate is zero. Discussing playing solitaire

competitively and single person games seems trickier for the advocates of meta-institutional concepts. You can build an argument against meta-institutional concepts based on this: Is the same subject going to call himself 'blasphemer'? In order to have the ordinary game and the meta-institutional game it seems that you need two parties viewing the same technical end in two different ways.

There is still a lot to unpack both to enhance the concept of technical end or to develop a less naïve theory of meta-institutional concepts. Whatever these developments may come to be, I hope it is now clear that chess won't be enough as a starting point.

References

Azzoni, G. 1988, *Il concetto di condizione nella teoria delle regole*, Pavia: Cedam.

De Filippi, P. & Wright, A. 2018, *Blockchain and the Law: The Rule of Code*, Harvard: Harvard University Press.

Feis, G. 2014, "Barare, copiare, doparsi, simulare", *Rivista internazionale di filosofia del diritto*, 99-103.

Feis, G. 2018, "Impossibility, Non-Existence and Constitutive Rules", *ARSP*, 104, 2, 284-97.

Guastini, R. 2012, "Manifesto di una filosofia analitica del diritto", *Rivista di filosofia del diritto*, 1, 1, 51-66.

Haas, W. 1953, "Value-Judgments", *Mind*, 62, 248, 512-17.

Karhulahti, V.M. 2013, "Puzzle Is Not a Game! Basic Structures of Challenge", *Proceedings of DiGRA* 2013: DeFragging Game Studies.

Leonardi, P. 1983, *Sulle regole*, Verona: Libreria Universitaria.

Lorini, G. 2012, "Can Constitutive Rules Create a Practice?", *Praxis Filosófica*, 34, 139-48.

Lorini, G. 2014, "Meta-institutional Concepts: A New Category for Social Ontology", *Rivista di Estetica*, 56, 2, 127-39.

Miller, D. 1981, "Constitutive Rules and Essential Rules", *Philosophical Studies*, 39, 183-97.

Moore, M. 2011, *Basics of Game Design*, New York: CRC Press.

Myers, D. 2003, *The Nature of Computer Games: Play as Semiosis*, New York: Peter Lang Publishing.

Roversi, C. 2007, *Pragmatica delle regole costitutive*, Bologna: Gedit.

Roversi, C. 2014, "Conceptualizing Institutions", *Phenomenology and the Cognitive Sciences*, 13, 1, 201-15.

Roversi, C. 2015, "Acting Within and Outside an Institution", *Methode*, 4, 191-212.

Sageng, J.R., Fossheim, H.J., and Larsen, T.M. (eds.) 2012, *The Philosophy of Computer Games*, New York: Springer.

Schwyzer, H. 1969, "Rules and Practices", *Philosophical Review*, 78, 451-67.

Torrengo, G. 2016, "Institutional Externalism", *Philosophy of the Social Sciences*, 47, 1, 67-85.

Wetzel, L. 2008, *Types and Tokens: An Essay on Universals*, Cambridge, MA: MIT Press.

Żełaniec, W. 2013, *Create to Rule: Essays on Constitutive Rules*, Milano: LED.

Constitutive Rules: The Manifest Image and the Deep Image

Maurizio Ferraris

University of Turin

Abstract

Social objects originate from constitutive rules. But there are two ways of explaining the relationship between them. I call them "Manifest Image" and "Deep Image". The former depends on Searle's interpretation of social reality and it is based on collective intentionality; the latter is the one I support and it is based on documentality. Indeed, recordings and documents are sufficient to explain how and why social world exists. There is no need to use such a vague notion, as that of collective intentionality, in order to give a useful account of society. Documents can do it better, especially with the help of the process called *emergence*, as the case of money clearly shows.

Keywords: Social world, Documentality, Intentionalism, Emergence.

1. Introduction

Constitutive rules are rules that constitute social objects. For instance, the rules of tennis constitute the game of tennis, and the rules of the Italian constitution constitute Italy as a state. But constitutive rules do not arise from nothing. Where do they come from? According to Searle (1995 and 2010), constitutive rules are an outcome of collective intentionality. However, Searle himself acknowledges that collective intentionality in turn needs to rest on something non-intentional, which he calls "the Background". Yet, as Rust (2009) pointed out, Searle finds it hard to explain what the Background really is and how it really works. Furthermore, it is debatable whether the Background limits itself to support collective intentionality or, instead, can produce social objects on its own (cf. Terrone and Tagliafico 2014). Finally, there might be other factors that ground social reality over and above collective intentionality and its Background (cf. Epstein 2015).

In this paper, I will argue that there is a layer of recordings and documents that grounds constitutive rules and therefore social objects at a deeper level than that of collective intentionality. I will state that recordings and documents are the cornerstone of the empirical background that warrants the production of social objects through constitutive rules. In this sense, I will propose to move from

Argumenta 4,1 (2018): 85-93
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.fer

a "Manifest image", i.e. a conception of social reality based on collective intentionality, to a "Deep image", i.e. a conception of social reality based on recordings and documents.

There is nothing more wrong than the idea posited by Vico—and by many after him—that nature is obscure because it is God's work, while society is transparent because it is man's work. First of all, it is unclear where nature ends and society begins; secondly, most of us surely find it hard to conceive of ourselves as the men who created society. In fact, our relationship with society is no different from our relationship with nature: that is, one of competence without understanding. If we throw a stone up in the air we move to dodge it even before learning the law of gravity, and we promise, we bet, and write wills based on vague, and often wrong, notions of law and society. If that is the case, there is nothing strange about the fact that we may often be surprised by social reality, which may reveal hidden aspects (after all, the surplus value was unknown to both capitalists and workers, as well as economists, until Marx discovered it).

The structure of society is not transparent to its members any more than the structure of consciousness is transparent to its subject. The fact that those things are the closest to us is anything but an advantage: it's rather a case of something being "a little too obvious", like the purloined letter of Poe's novel, which for that very reason escapes our observation and conscience. That is why, to answer the question (paradigmatic for social ontology) of what constitutes money, i.e. gives the value of money to a piece of paper, one cannot point to some evidence, but rather has to solve a riddle—or at least try to do so. One needs to rip a veil that hides the whole sphere of social normativity: What is a constitutive rule, i.e. what is the secular sacrament that transforms a human being into a doctor (authorized to cure), a licence holder (authorized to drive), a recipient of a call-up paper or a payment order (required to show up to the barracks or to pay)? . Interestingly, in all these cases there is indeed a piece of paper involved—although, of course, the paper might be replaced with plastic, metal (still in use for money), tattoos, distinctive and picturesque signs of all sorts, or simply recordings on a computer or a mobile phone. To solve a riddle, one must first of all understand the terms involved. Thus, in order to figure out what grounds constitutive rules, I will highlight their dependence on recordings and documents. I will do so by analysing two possible conceptions of social reality, which I will call the Manifest Image and the Deep Image, respectively.

2. Manifest Image

"Manifest Image" does not mean "false image. In fact, it is tempting to see society and its objects, including money, as the outcome of our intentionality (money has value because we think it does, laws apply because we think they do). From this perspective, the most direct and intuitive way to explain the functioning of money and social reality as a whole is intentionalism, whose most illustrious interpreter is indeed John Searle.[1] For Searle, the constitutive rule of social objects is "X counts as Y in C": X (the physical object, e.g. a piece of paper) counts as Y (social object, e.g. a banknote) in context C, because of collective intentionality. That is what I will call the Manifest image. Searle's intentionalist perspective has a twofold structure. Its first element is the claim that social reali-

[1] Searle 1995 and 2010.

ty is constructed by us. In fact, when one reflects on our relationship with money, titles, works of art, etc., one might feel that these objects are socially what they are because we collectively decide that they are, indeed, money, titles, works of art. The second element is the thesis that this "us" manifests itself as collective intentionality (a close relative of Rousseau's general will, of Montesquieu's spirit of the laws, and, after all, of the spirit in the Christian and Hegelian sense). In this sense, the manifest image may meet some fundamental intuitions. For instance, when we pay, we may have the impression that both us and the recipients of our money share the conviction that money has value because we believe so.

However, the manifest image leaves a few things unexplained. First of all, it is unclear what is meant by "collective intentionality": a vague notion that seems to not only cover obvious cases of sharing actions ("how about we take a walk?" "let's have a coffee") but also chimerical constructs like the general will. It is also unclear what its spatial-temporal location would be, provided there is one. Instead, it is very clear what individual intentionality may be—something that can be shared and coordinated with others on the basis of a document. All in all, collective intentionality is but a legal and philosophical fiction comparable to the generalizability test of the Kantian moral imperative: when the court issues a ruling "in the name of the Italian people" it is acting as the representative of a collective intentionality? Of course not: it simply means that the decision is not arbitrary, and is taken in compliance with the law. Likewise, expressions such as "the Court ruled" do not express collective intentionality, but simply a decision (taken by majority or unanimously), that is, the numerical predominance of individual intentions.

Secondly, the intentionalist perspective does not account for negative entities, such as debts—it is difficult to find a physical X corresponding to the negative social Y.[2] The same difficulty applies to electronic money. If I pay with my cell phone, is the physical X the phone? If it is then the same object has two prices: a sales price and a variable price, which manifests itself through its purchasing potential, which could hardly be considered a property of the phone as a physical X.

Finally, collective intentionality interprets society in terms of harmony and consensus rather than in terms of conflict, disagreement, contradiction: and yet the latter is the way in which social reality has always appeared to us, from the *Iliad* to today. Society, as well as normativity (laws, obligations, institutions, rules, prohibitions), does manifest itself not only in consensus but also in conflict. Indeed, norms are mainly perceived when they are in conflict with our instincts and our immediate dispositions, clashing with what we would like to do. And collective intentionality only explains some situations in which normativity is weak: a walk or a picnic, not a board of directors, a high command or a court. The same sharing of collective intentions that seems to unite the members of a football team or the musicians of an orchestra is the result of a document-based normativity: in the first case the coach presenting the game schemes and the constraints imposed by the rules of football, in the second the director (whose presence would be completely useless, like that of coaches and generals, if there really was collective intentionality) and the sheet music.

[2] Searle and Smith 2003.

It is not surprising that the manifest image should face the same difficulties as the social contract in politics, and as dualism in the theory of the mind. These difficulties become particularly notable in the case of money, whose structure would be divided into a spiritual part that is in us (the value attributed to money), and into a material and accidental part that is outside of us. Intentionality would be acting as a collective pineal gland, called to link the immaterial (the value, the social object) with the material (the piece of paper, the physical object). In short, the manifest image undoubtedly explains certain social acts—there is no question that a parliament constructs something when promulgating a law (even though it is worth noting that the form of the law and its context are already there, and therefore it isn't an absolute construction). However, this intentionalist explanation, if applied to the whole social reality, seems to be no less mythological than an explanation of morality for which the ten commandments are actually the manifestation of God's will to Moses.

In fact, there are many empirical circumstances disproving the intentionalist explanation. First, the obvious difficulty that it is impossible to determine when and how the "invention" of money actually took place. Secondly, the even greater difficulty of clarifying the nature of collective intentionality, which carries the burden of proof. Thirdly, and most importantly, the intentionalist explanation makes money a fragile invention which could be rejected at any time by the mere end of consensus. But, in fact, this is false. The reduction (not disappearance) of monetization in the Middle Ages was not the result of consensual agreement, but depended on the rarity of coins, which were no longer able to cover the amount of real exchanges. Also, the enemy's money is still valid in times of war (as Wittgenstein's father well knew, investing his capital in the titles of the Entente).

Also, whatever collective intentionality wants, or does not want—and provided such a unitary feeling exists—when a state prints too many banknotes people do not become richer (as they should if the value of money depended on collective intentionality) but poorer, and money loses value. Indeed, although everyone agrees that 1,000 marks is a lot of money, why is it that they can suddenly be worth nothing, and one has to switch to banknotes like 100,000, or 1,000,000? It would seem that we are dealing with a collectively masochistic intentionality, rather than with a collective intentionality. Lastly, if collective intentionality determined the value of money, it would be impossible to explain phenomena such as financial crises. Neither the latter nor natural phenomena can be controlled. Yet, there is a single and significant difference, which relates to the their different deep structure: namely, that the disappearance of collective memory and documents would put an end—albeit dramatically, meaning the end of civilization—to a financial crisis, but it would not be able to stop the rain, nor to question the law of gravitation.

3. Deep Structure

To move to the deep structure, I invite you to act like the fool of the famous Chinese proverb: do not look at the moon, but at your finger. In this case, do not look at the mind and the wonderful representations it contains, let alone at collective intentionality (provided you can find it), but rather look at the notes you have in your wallet or the change tinkling in your pocket. It is worth noting that in Searle's intentionalist formulation of the constitutive rule, "counts as"

can easily be translated into "stands for," which is the character of the symbolic relationship. From this perspective, social objects created in this way are all symbols, but of what? Of the Fort Knox Gold? Of course not. Of the ideas you have in your head? Neither, as you can have a lot of money but a few ideas, and vice versa. Of what, then? What if they were not symbols, but real objects, that, *far from representing states of mind, are actually able to produce them?*

There are two main theses following from the idea of the deep structure. The first is that collective intentionality does not exist; instead, there is an often-conflicting social interaction that is made possible by the use of documents (both in the strict and in the broad sense: institutions, rituals, transmitted behaviours) that coordinate individual actions and intentions. The second is that money exerts its prestige on individual intentions without any intervention of collective intentionality—for the very good reason that a non-existent entity has no causal value—and based on the sole force of what I call "documentality".[3] By this term I do not mean the sphere of the intentions that exist in our mind, but that of social recordings, from the promise onwards, which exist both out of our minds—in archives, wallets, cell phones—and in them, but as external elements: think, for example, of our memory of a word given, which is no longer entirely ours (unlike what happens to so many other thoughts that belong to us and only to us).

The documentalist explanation—that is, the deep structure—is structured in two theses: the first is that documentality is the necessary condition of social reality, which cannot exist without documents; the second is that documentality is the sufficient condition of social reality: if there are documents, along with beings biologically identical to us (in particular, endowed with sensibility and memory), there is everything that makes up social reality, including individual and collective intentionality. The documentalist explanation, just like the intentionalist one, depends on the theory of speech acts[4]: there are acts that do not just describe or prescribe, but actually construct objects: a marriage, a debt, a holiday, a war. Only, instead of taking the consistency of the act (its ontological status) to depend on the physical objects that are transformed into social objects, it posits that the transformation of the act into an object depends on recording, according to the formula: *Object = Recorded Act*. The social object is the result of a social act (involving at least two people, or a person and a delegated machine) that has the characteristic of being recorded—thereby acquiring the permanence typical of objectivity—on any physical medium. This, among other things, easily accounts for negative entities: debts are noted in the column of giving, just as credits are in the column of having.

The crucial role of recording is very clear in the case of money. I have a note in my hand, and I can pay the bill at the restaurant. I can do so also with a debit card, with a cell phone, or photographing a barcode on the bill. What do these operations have in common? The fact that there are recordings—analogue or digital codes on my account, analogue or digital memories in my pockets—like paper tickets, plastic cards, or even a phone, which can do a lot of things precisely because it has a lot of memory, which results in a great computational capacity. So, money is a form of recording. In fact, all money can be traced back to this origin and function—and anything that can accomplish this function can act as money. There is *no change* in terms of the nature of money occurring be-

[3] Cf. Ferraris 2005, 2009, 2014.
[4] Austin 1962.

tween a note, a card and pure memory: what occurs is rather a revelation of what money really is, namely the recording of a numerical value that, through recording itself, acquires economic value. Certainly, to be a *valid* recording has to reflect the economic value of the merchandise transferred or service rendered, not just *any arbitrary* recording. Yet, this just shows that a recording cannot function on its own. Rather, it needs to be included into a wider network of recordings, including those in people's minds (cf. Terrone 2014a, 2014b).

If we understand that the essence of money is recording (according to the rule Object = Recorded Act, which is particularly clear with banknotes, where the act is a relation between the central bank and the anonymous holder of the money), we can also understand why, before banknotes, people used coins (perhaps of gold, a material that does not rust), or shells, or salt sacks: all discrete portions that can be counted, generating an archive; that can be subdivided, facilitating payments; that can be kept in a limited space (for this reason only, by the way, coins are better than salt bags). This shows that the constitutive rule that makes money what it is requires much more than collective intentionality. In fact, constitutive rules require a background of empirical conditions, which come from the role that recording plays in the production of these rules. Constitutive rules are not just a matter of intentionality but also—indeed, first of all—a matter of matter. In this sense, the power behind money, recording (the genetic principle of the form), is the principle underlying social normativity as a whole. According to what we have just said, documentality is the principle of responsibility, which in turn originates normativity—indeed, and more exactly, it is what Montaigne and Pascal called "the mystical foundation of authority". In this sense, the essence of money is manifested in the bitcoin, and retrospectively the bitcoin makes the value of the gold coin, of gold, and of salt real. The digital currency, in fact, is nothing but the memory trace of a transaction, a pure document that has no external rooting, if not a secure and public record (the blockchain) that registers the transaction and acts as its guarantor.

This, like other empirical facts, proves the validity of documentalism. Society cannot do without inscriptions and recordings, archives and documents, and without the arche-technology of writing, which is the prototypical form of recording. Moreover, without recording there would not and could not be legal institutions, obligations, guarantees and rights. So, justice would never have been fully realized, as it is intrinsically social. Documents do not only act as regulators in the economy and in the legal sphere, but are the producers of values, norms, cultures, conflicts, up to determining (through education and imitation) individual intentionality and allowing (through sharing) for collective intentionality. Despite appearances, it is the document that creates the value, not the value that produces the document: gold is not worth it because it is gold, but because it has characteristics (the same ones that make it a useful metal in jewellery) that make it a durable and malleable document medium. Documents, in this sense, are the cornerstone of the empirical background of constitutive rules.

4. Pentecost or Emergence

The contraposition between intentionalism and documentalism implies a metaphysical problem. Considering (collective and individual) intentionality as a primitive leads us to embrace what I call "Pentecostal meaning": that is, postulating the existence of a meaning previous to and independent of the forms in

which it is expressed and of the ways in which it is imprinted—that is, the psychological and social equivalent of Cartesian dualism. This view involves pattern of this kind: in the mind there are meanings that are expressed through words, which in turn are represented in writing. So, meaning might exist even if unexpressed, and, most importantly, meaning has no genesis: it has always existed or has fallen from the sky. This model is found in most theories of man and of society. For instance, it is often postulated that there is an in-itself, human nature, which is alienated by external conditions, usually associated with technology, and which must be restored through a return to human nature as it really and naturally is. In such a theory of society, the origin of the social world is placed precisely in collective intentionality, which manifests itself through a contract from which society originates.

Indeed, from the intentionalist perspective, money and its normative power are a variation of the social contract: it is agreed to give value to a piece of paper, or gold, a shell or a salt sack, just as it is agreed to regulate society in a certain way. The counterpart of this approach, in theory of the mind, is the postulate of a *res cogitans*, distinct and independent from the *res extensa*. All of this is based on a precise topology: meaning, spirit, idea, and consciousness are inside; signifier, letter, expression, and action are outside. Conversely, the documentalist explanation calls for an emerging meaning (meaning comes from act and recording) instead of a Pentecostal one (meaning precedes act and recording). If Pentecostal meaning is conceived as independent and anterior to expression and recording, emergent meaning, on the contrary, recognizes its dependence on both, and proposes a Copernican revolution that consists in overturning the traditional structure and conceiving intentionality (the spirit, the idea, the will, and the purpose) as successive and derivative, rather than as prior and foundational, compared to the forms of fixation (the letter, the expression, the norm). More radically, the deep structure shows that *documentality is a condition of intentionality*. Surely, written symbols need minds in order to acquire meaning, but minds themselves require some forms of fixation, which in the contemporary debate in cognitive sciences and the philosophy of mind have been characterized in terms of mental files (cf. Recanati 2012 and Terrone 2017 for an application to social ontology). From this perspective, the claim that intentionality requires documentality can interpreted as the claim that intentionality requires the deployment of mental files, which are the mental counterparts of documents.

This change of perspective overcomes the difficulties raised by the manifest image, and in particular it answers the question why, if social reality is constructed, it is so difficult to change it. The answer is precisely that the manifest image hides an essential point: *the fact that social objects are constructed does not mean that social reality is constructed*. Like money, society is not constructed, but emerges. Above all, society is not just a human fact.[5] Society is not simply composed of humans, but includes dimensions other than human (animals, for example), or superhuman (myths, which are constitutive elements of the social world). Such dimensions are the structures that make us human. If it is difficult to imagine non-human animals investing in the stock market (but not exchanging banknotes!), it is even more difficult to imagine that our forms of social organization (dominance structures, elementary kinship relationships, taboos) have no relation of continuity with our animal past. Likewise, it is difficult to

[5] For a criticism of anthropocentric social ontology cf. Epstein 2015.

imagine a human social activity that would not be decisively conditioned by its technical forms of realization. And recording, as I have argued above, is the most basic technical form of realization that grounds social reality.

So, money, very simply, is a document like any other: it's like a passport, for example, and shares its complicated decorations and characteristic colours (blue for Americans, red for Europeans, green for Arabs, as far as passports are concerned, that is). With a passport a state authorizes a citizen to expatriate (so it was originally) and with a banknote it authorizes her to buy. Since the citizens who want to buy are far more numerous than those who want to expatriate, there are more banknotes than passports. And since money goes from hand to hand, banknotes are "on bearer", as exchanges are made quickly, and possibly by illiterate people—in most countries (albeit with the significant exception of the United States) banknotes have different size and colour, so that money could be defined as the documents of those who cannot read. Here is another important sense in which documentality contributes to the empirical background of constitutive rules, as far as the rule that turns a piece of paper into money exploits the empirical features of this very piece of paper. In addition, both with passports and with banknotes, the state does not invent anything new, but merely gives a paper form to ways of fixating acts and quantifying value that originated in our animal past and whose evolution coincides with the evolution of human cultures.

Ultimately, if one focuses on the Manifest image of social reality, one may have the impression that collective intentionality is almost almighty as far as it can create social objects at will by simply stating the corresponding constitutive rules. Yet, if one shifts one's attention onto the Deep image, as I have tried to do in this paper, one can acknowledge that the creation of social objects through constitutive rules has much deeper roots, which on closer inspection reveal to be documents and recordings.

References

Austin, J.L. 1962, *How to Do Things with Words*, Cambridge, MA: Harvard University Press.

Epstein, B. 2015, *The Ant Trap: Rebuilding the Foundations of the Social Sciences*, Oxford: Oxford University Press.

Ferraris, M. 2005, *Dove sei? Ontologia del telefonino*, Milano: Bompiani; Engl. transl. *Where are you? Ontology of the Mobile Phone*, New York: Fordham University Press, 2014.

Ferraris, M. 2009, *Documentalità. Perché è necessario lasciare tracce*, Roma-Bari: Laterza; Engl. transl. *Documentality. Why it is necessary to Leave Traces*, New York: Fordham University Press, 2012.

Ferraris, M. 2014, *Mobilitazione totale*, Roma-Bari: Laterza.

Ferraris, M. 2016, *Emergenza*, Torino: Einaudi.

Recanati, F. 2012, *Mental Files*, Oxford: Oxford University Press.

Rust, J. 2009, *John Searle*, New York: Continuum.

Searle, J.R. 1995, *The Construction of Social Reality*, New York: The Free Press.

Searle, J.R. 2010, *Making the Social World: The Structure of Human Civilization*, Oxford: Oxford University Press.

Searle, J.R. and Smith, B. 2003, "The Construction of Social Reality: An Exchange", *American Journal of Economics and Sociology*, 6, 2, 285-309.

Terrone, E. and Tagliafico, D. 2014, "Normativity of the Background: A Contextualist Account of Social Facts", in Gallotti, M. and Michael, J. (eds.), *Perspectives on Social Ontology and Social Cognition, Series: Studies in the Philosophy of Sociality*, Dordrecht: Springer, 69-86.

Terrone, E. 2014a, "Traces, Documents, and the Puzzle of 'Permanent Acts'", *The Monist*, 97, 2, 161-78.

Terrone, E. 2014b, "Che cosa esiste fuori dal testo?", *Rivista di Estetica*, 57, 101-20.

Terrone, E. 2017, "The Band of Theseus. Social Individuals and Mental Files", *Philosophy of the Social Sciences*, 47, 4-5, 287-310.

Constitutive Rules, Normativity, and A Priori Truth

Terry F. Godlove, Jr.

Hofstra University

Abstract

This paper develops an argument which seems to yield a set of a priori rules—rules which are constitutive of, but not normative for, thought and experience. I contrast the resulting Kantian sense of a priori truth as independent of all experience, because presupposed by it, with the use Searle makes of a priori truth by stipulation or definition. By focusing on the a priori rules of thought and experience we can make good on the sense of constitutivity that Searle had in mind in his early work. By virtue of their apriority, the Kantian rules are able to do what the constitutive rules of football and chess cannot: they are able to define the nature of the activity they govern, namely, thinking or cognizing that thus-and-so. They tell us, independently of their cultural or social context, what kind of activity results from our compliance with them.

Keywords: Constitutive, Rules, A Priori, Normative, Experience, Kant, Searle.

It is now almost fifty years since John Searle began calling attention to the relationship between "constitutive" rules and forms of behavior. He writes, in 1969, that "constitutive rules […] create or define new forms of behaviour. The rules of football or chess, for example, do not merely regulate playing football or chess, but as it were they create the very possibility of playing such games" (Searle 1969: 33). Over the years, Searle's discussion of constitutive rules has been influential, and rightly so, in philosophy of language, social ontology, and elsewhere. But from early on, it has been clear that Searle's analysis requires supplementation and revision. For example, Hubert Schwyzer, and more recently, Giuseppe Lorini, have pointed out that, while activities like football and chess are played in accordance with rules, they cannot, as per Searle's original claim, be constituted or defined in terms of those rules (Schwyzer 1969; Lorini 2012, 2014; see also Bierman 1972). To understand what makes possible a game of football or chess requires casting a wider cultural net than Searle has in hand. People can conform to the rules of chess—they can move the pieces according to the rules—and yet be doing very different things. Are they moving the pieces with the aim of winning, or do they have purely aesthetic goals in mind? Or perhaps for one or both of them it is a purely instructional exercise.

Argumenta 4,1 (2018): 95-109
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.god

In what follows I will be writing in this same spirit, that is, as supplementing and correcting Searle's work on constitutive rules. But whereas Schwyzer, Lorini and others have urged greater attention to the material context of constitutive rules—to the social, cultural, and institutional contexts in which they are embedded—I will be applying corrective pressure from the opposite direction. I will be arguing, from a generally Kantian point of view, that Searle-style constitutive rules take place in a formal context as well—in a context of higher-order rules which are constitutive, not of this or that form of behavior, but of any form of cognitive engagement with the world; rules the following of which make experience possible. The prospect of further, higher-order constitutive rules, raises the threat of regress. But instead of constitutive rules "all the way down," in prospect is an ontology general enough to encompass the physical and the social world—all the world there is. Two themes will predominate. First, I will be urging that we shift our attention from Searle's emphasis on analytic truth to a priori truth; this, I will argue, is the more fundamental notion. Second, I will be keeping in view the question in what sense, if any, these higher-order constitutive rules have a normative dimension.

1. Normativity

Since, for Searle, rules become constitutive through stipulation or definition it is only natural that he should emphasize analytic truth.

> what the 'rule' seems to offer is part of a definition of "checkmate" or "touchdown". That, for example, a checkmate in chess is achieved in such and such a way can appear now as a rule, now as an analytic truth based on the meaning of "checkmate in chess." That such statements can be construed as analytic is a clue to the fact that the rule in question is a constitutive one (Searle 1969: 34).

As I understand it, the idea is that a rule is constitutive of a certain practice when it is attached to the relevant general concept by way of the concept's compositional structure. Searle is apparently helping himself to a traditional and rather powerful way of thinking about general concepts, one that goes back at least to Boethius. Concepts can be represented as conjunctions of two or more partial concepts; for Boethius, "homo" stands under "animal rationale" by way of "mortale" but not "immortale" (see De Jong 1995). An affirmative judgment is analytic just in case the subject concept can be analyzed as a conjunction of concepts one of which is the predicate concept. Thus, we might legislate that to checkmate someone is to put his or her king in a check from which it cannot escape. Or we might create a concept, "saying," which will refer to the act of expressing a proposition by means of uttering a declarative sentence (Cappelen 2011). Having stipulated the constituents or parts of the subject term, I cannot then think the concept "checkmate" without also implicating "king," "escape," etc.—and I cannot claim the right to checkmate someone while leaving his or her king an escape route. Nor can I think "saying" without implicating "utter," "declarative," "sentence," etc.—and I cannot claim the right to say something except in the declarative voice. Having arranged in advance for their inclusion in the subject terms, it is no surprise that these concepts and rules emerge in the course of their analysis.

This basic account of how constitutive rules emerge from the compositional structure of general concepts has much to recommend it. Difficulties set in when

we ask too much of it, as, for example, when we ask it to explain or rationalize culturally imbedded human activities and practices. Thus, suppose I reach out to move one of my bishops. Under the right circumstances, my behavior will be a move in a game of chess and will have been made possible by constitutive rules in Searle's sense—among them, that it is legitimate to move a bishop only along a diagonal of its own color. The burden of Schwyzer's and Lorini's critique is that—as illustrated in my first paragraph—explicating "the right circumstances" will require us to draw on material resources well beyond the constitutive rules of chess.

I think the force of this critique has to be admitted; constitutive rules cannot create or define a new (or any) culturally embedded activity or practice. However, the model of constitutive rules as reflecting the compositional structure of general concepts retains its power and is of independent interest. Searle is right that we would not ordinarily say, of someone who moves her bishop other than along a diagonal of its own color, that she is playing chess. Ordinarily, we would say: If you want to play chess, then you must conform to this rule. Here "must" has all and only the force of our signing on to use the general concepts at issue in the agreed-upon way.

To begin to expose what I am calling the formal context of Searlean constitutive rules let us now ask whether they have normative authority over the actions they govern. And let us accept the criteria Clinton Tolley sets out for normativity. In order to have normative force over a certain practice

> (1) The "subjects" of the law—those beings which are governed by, or subjected to, the law—must both be able to succeed *and be able to fail* to act (or be) in accordance with the law.
>
> (2) The *subjects* of a norm *must retain their identity* as beings that are subjected to this specific sort of law regardless of their (actual) accord with it. This latter condition is important, as it implies that evaluative ascriptions in light of norms (e.g., *x* as "in" or "out of accord") institute a division *within* some otherwise well-defined class.
>
> (3) The *laws must retain their validity* or bindingness over their subjects regardless of the (lack of) *actual* adherence to the norms by their subjects—though, to be sure, there must be the *possibility* of such adherence (Tolley 2006: 375, original emphasis).

Tolley points out that, by this standard, traffic laws can be considered normative for the activity of driving. Drivers can succeed or fail at them without jeopardizing their status as drivers, and the laws retain their prescriptive force even if, any a given time, no one is following them. The same cannot be said for the rules of chess. Failure to observe the rules puts me in violation of Tolley's first two conditions: Moving my bishop except along its diagonal does not count as a move in the game, nor, except by conforming to the rule, do I retain my identity as a chess player. By contrast, I can drive through a red traffic light—whether legitimately (to make way for an ambulance) or not (on a dare)—and, in doing so, I retain my identity as a driver. So, at least on Tolley's criteria, the rules of chess are constitutive but not normative for the activity they govern, while the rules of the road are normative but not constitutive.

We will presently take up the suggestion that certain rules could be constitutive for a domain of cognitive activity and normative for it as well. Prima facie, this is an odd-sounding prospect, for it envisages a rule that permits its own

violation while at the same time it emerges from the compositional structure of the general concept at issue. Before coming to this question, let us consider the rules of chess and the rules of the road together and ask whether there are rules with greater generality that govern both—whether constitutive, normative or both.

A ready candidate—one about which philosophers have had much to say—is the principle of non-contradiction: not both p and not-p, or

(4) ~(p & ~p).

Suppose I claim the right to move my bishop and not move my bishop to the same square or to both stop and not stop at the red light. A moment's thought shows that the oddness in these claims has not to do with the physical movements—has not to do with my hand both reaching and not reaching for the piece or with my foot both pressing and not pressing the brake pedal. The difficulty with a rule that commands me to, "Move and not-move" or "Press and not-press" comes prior to any question of physical movement. I am unable to form the thought of moving and not moving, pressing and not pressing—so the question whether such putative rules are constitutive or normative or both does not arise.

What about the principle of non-contradiction itself? The invited conclusion—that no conduct-governing rule can be self-contradictory—apparently applies beyond the domains of chess and driving. The invitation is to conclude that it applies generally to what can be asserted, and so to any domain or activity whose form is that of thinking or cognizing that thus-and-so.

2. Logic

We might be tempted to bestow both constitutive and normative status to a rule of such apparent importance—constitutive because following it makes possible the associated mental activity; normative because the advice to avoid explicit self-contradiction would seem to be a worthwhile bit of practical guidance for our thinking and acting. A well-known remark attributed to Kant suggests this approach. In the so-called *Jäsche Logic* we read:

> In logic […] the question is not about […] how we do think, but how we ought to think. In logic we do not want to know how the understanding is and does think and how it has previously proceeded in thought, but rather how it ought to (*sollen*) proceed in thought (Kant 2004: 529).

We find similar remarks scattered throughout Kant's minor writings, many seeming to claim that logic has normative authority for reasoning—what Huaping Lu-Adler calls the "*sollen* claim" (Lu-Adler 2017).[1] A renewed focus on these remarks has sparked controversy in the recent literature, with Tolley urging their rejection in favor the view Kant expresses in the *Critique of Pure Reason* (Tolley 2006; see also Mosser 2008: 43 ff, and Pollok 2017: 8).[2] Kant there says that "general" or "pure" logic "contains the absolutely necessary rules of thinking, without

[1] For a survey of the passages in Kant's minor writings see, especially, sections 3.2 and 3.3 of Lu-Adler's paper.
[2] I return to Pollok's views, below, sections 6 and 7.

which no use of the understanding takes place" (B76).[3] Under this heading Kant clearly means to include the principle of non-contradiction; it is, he says, "the highest principle of all analytic judgments" (A150/B189)—in which capacity, let us note well, it makes possible constitutive judgments as Searle conceives them: Having stipulated to the contents of a general concept, we can then violate the paired rules of conduct only on pain of self-contradiction. I cannot both agree that a checkmate ends the game and reserve the right to keep playing the game after having been checkmated.

In an earlier paper (cited by Tolley), Manley Thompson emphasizes that the issue at stake is not that of inconceivability:

> The point [is] not that we cannot conceive of or imagine ourselves meaning something in a situation where we rejected the principle of contradiction. It [is] our inability actually to mean (not just to conceive of ourselves as meaning) anything in such a situation (Thompson 1981: 464).[4]

If we cannot help but presuppose the principle of contradiction, then we may say (with Thompson) that the legitimacy of following this rule in our thinking arises from the nature of thought itself—surely Kant's point at B76. (Lu-Adler points out that the apparent tension in Kant's views on this point can be resolved by following his distinction between pure and applied logic, as, for example, at A53-54/B77-78 [Lu-Adler 2017: sects 2.4 ff].) This would then be the consideration grounding Tolley's three criteria of normativity, each of which can now be seen as parasitic on the more fundamental claim that a rule can have normative status for thinking creatures only if conformity to it is not required by all of our thinking or cognizing that thus-and-so.

Thus, if we hold to Tolley's criteria of normativity, the principle of non-contradiction and whatever else depends upon it cannot be accorded normative status. If the principle of non-contradiction is constitutive for the activity of asserting or cognizing that thus-and-so—if conforming to it makes that activity possible—then what sense does it make to say we ought to conform? In so saying we again run afoul of the first two of Tolley's criteria: We cannot fail to conform to it and still judge that thus-and-so. And we cannot both fail to conform and retain our identity as cognizers—that is, as subjects who experience the world as thus-as-so (as subjects who see, for example, that the bishop is there and not not-there and that the light is red and not not-red). If we want the principle of non-contradiction to be constitutive for some domain or activity then it appears we cannot at the same time give it normative status.

To look ahead for a moment: Our interest is going to be in social ontology—in the rules that form the conceptual background for or that make possible social reality. In a Kantian spirit, what I want to suggest so far is that, in the principle of non-contradiction, we have a rule constitutive for any object of possible experience; for any object, physical or social, that we could be in a position to cognize. A rule that, at the same time, cannot be a norm for creatures like us.

[3] References to the *Critique of Pure Reason* follow the usual A (1781) and B (1787) format, and will henceforth appear in the text. I generally follow the Paul Guyer and Allen Wood translation (Kant 1998).

[4] I have benefitted from Mosser's discussion of Thompson's paper (Mosser 2008: 45 ff).

3. Singular Reference

We will turn, in section 6, to an objection to this line of thought, but first let us ask whether the principle of non-contradiction stands alone or whether there are other constitutive rules which are, as it were, too deep to be norms.

In reaching out to move my bishop I am affected by it in sensation. I cannot create the bishop merely by thinking of it; in Kantian terms I have a discursive or sensible rather than a divine intellect. To cognize this bishop I must be in a passive state with respect to it. Nor can I substitute for the affection (for the sensation) the concept or general term "bishop". No matter how finely I slice the general concept (the black Staunton-style one with the chip off the corner, etc.) my description could apply to more than one object. Rather, my connection to this bishop is immediate and singular—a Kantian *empirische Anchauung*.

In reacting, "bishop," to what strikes me in sensation, I locate the object at some distance from me. (I also locate myself in time, but I cannot pursue the status of temporal rules in this paper.) That is, I follow a rule which says,

(5) Locate objects at some distance from where you find yourself.

This is of course not a rule of chess. But it is a rule conformity to which governs the activity of chess-playing. In fact, it appears to be constitutive of any activity involving a judgment or judgments that objects outside me are thus-and-so—constitutive, as above, in the sense that following it makes possible the activity in question. Thus, we cannot suggest, in a pragmatic spirit, that conformity to this rule makes for enhanced success at chess-playing or for running fewer stop signs. I presuppose my conformity to this rule in locating the bishop or the stop sign at a distance from me in engaging in either activity.

The case of blindfold chess is instructive. David Shenk reports the great German player Siegbert Tarrasch as saying, "Some part of every chess game is played blindfold" (Shenk 2007: 126). That is, in contemplating my next move, I project the bishop moving on a line that I construct in imagination. For the purposes of this activity of imaginative construction, strictly considered, it does not matter whether I am literally blindfolded. What matters is that space is given to me as a whole, as a "horizon" into which I introduce limitations: Whether blindfolded or in contemplation, in order to consider where to move my bishop I must draw the line (a diagonal) in thought—and I know I can keep adding to that line without limit beyond the edge of the chess board. I draw a diagonal line for the bishop, or I go forward two units and turn 90° one unit for the knight, etc. These moves are not the manipulation of general concepts. They are glimpses of the future position of the piece that is now literally affecting me in sensation—glimpses afforded not by my magically travelling forward in time but by my shifting attention one unit at time, imagining the piece at incrementally different distances from me, that is, imagining it as affecting me at incrementally different distances. Space, on this Kantian view, is not itself an object (it does not affect me), nor is it a relation between objects (rather, I presuppose it in locating objects in relation to myself). It is the mode through which I am affected by objects outside me—a form of intuition.

We may now ask whether (5) is constitutive in the same sense as (4), the principle of non-contradiction. We saw that the latter is constitutive for the activity of thinking, and so for any cognitive commerce with objects of possible experience (commerce that requires judging of objects that they are thus-and-so). It is then, a priori, in a stronger sense than are Searle's constitutive rules. Whereas Searle gets constitutivity through stipulation or convention, our conformity to the

principle of non-contradiction is required if we are to think at all; its legitimacy is independent of experience, a priori, in that radical sense. By contrast, the rule requiring me to locate objects of cognition at some distance from me is not constitutive for thought *per se*; I can quite well think about objects in n-dimensional spaces—but it is constitutive for the cognition of objects of possible experience, e.g., bishops, knights, rooks, etc. For where, other than at some distance from me, would such putative objects be located? The question cannot be answered by appealing to what we can think without contradiction. It can only be answered by showing where, in relation to me, the chess piece (the one I wish to move) can be located except in three spatial dimensions. Rule (5), then, is a priori and yet synthetic—the combination whose provenance motivates Kant's Critical project.

In the terms I am developing, rule (5) presents us with a second case of background constitutivity, this time tied not to thought but to experience. Is (5) also a norm? When checked against Tolley's criteria, we can see that it is not. I am not free to play chess or drive my car and at the same time to fail to locate objects at some distance from where I find myself. I can think about playing or driving in n-dimensional space, but I cannot actually do so and at the same time retain my identity as a chess player or a driver. Once again, we have here to do with a constitutive rule with no normative dimension.

4. Categories

Let us now take (4) and (5) together—thought and experience. Arthur Melnick has proposed as the canonical form of singular reference

(6) "Take *n* steps Circumscribe Be affected React φ" (Melnick 1985: 48).

Melnick intends this to capture the prescription for sensing or obtaining sensation together with that of reacting (appropriately or not) to what strikes me in sensation. Besides reflecting the constitutivity of space as a form of affection, Melnick points out that this schema contains an additional element constitutive of singular reference. In reacting "bishop" I must employ what Kant calls the categories of magnitude: unity, plurality, and totality. Simply to cognize the bishop I must take its round base, its curved body, and its pointed top—a plurality—as a single figure, as a unity. And in contemplating moving it on a diagonal across the board I must generate the line in imagination. But, since a line is an infinite plurality of points, I am again taking a plurality as a unity—this time in the knowledge that I am proceeding toward, but will never attain, a totality of points.[5] The details are not to the point. What matters is that the constitutivity of space for experience brings along the concept of magnitude and its associated categories.

What about the concepts of unity, plurality, and totality—can my reliance on these categories be characterized in normative terms? Apparently not. As I look over the chess board, am I free to see the pieces (to play the game) without calling on the concept of plurality? As I consider whether to sacrifice one of my pair of bishops, can I do so without employing the concept of unity? And what about the line between my king and the threatening rook? I cannot construct a line in imagination without implicating the totality of points that make it up. Constitutive background rules, which, along with (6), are apparently devoid of normative significance.

[5] The argument of the Axioms of Intuition, A162/B202-A166/B207.

What we have so far is merely the form of singular reference—determination enough only to yield spatio-temporal somethings. By themselves, spatial intuition and its determinants, unity, plurality, and totality, are, as Melnick puts it, "ontologically neutral in that they do not settle questions of identity" (Melnick 1985: 55; see also Pollok 2017: 222-23). After Quine, the objects at issue could be bishop-stages, or bishop-histories, or undetached bishop-parts, etc. To cognize chess pieces or stop lights as enduring through time, as undergoing changes, and as causally interacting with other objects we must implicate the dynamical categories—those that deal not with magnitude but with existence and necessary connection. To demonstrate that these concepts are also constitutive of experience Kant invented a new argument form, a transcendental deduction—an argument to the effect that either I employ these categories or I do not experience at all.

My remarks on the form of singular reference have been brief; as concerns the dynamical categories I can have even less to say. But I hope I have put enough on the table to make out the fundamental difference between Searle's and Kant's interest in constitutivity. As indicated by the quote from Searle with which we began, his interest is in constitutivity by stipulation or convention. As we saw in section 1, he claims the rules of chess create "the very possibility" of playing the game because they follow from the sub-concepts that, taken together, form the parts of the general concept "chess"—hence Searle's emphasis on analyticity, and hence his vulnerability to such rejoinders as the one from Schwyzer and Lorini. As Searle has always emphasized, we are free to revise the composition of our concepts, and, so, the rules governing the paired behavior. Without this element of choice the question of deviant rules would not arise: thus, we could (and do routinely) decide to allow each chess player an agreed-upon amount of time in which to complete the game, or we might require one or both players to wear a blindfold. This is not to say that Searle's constitutive rules are not, in a perfectly good sense of the term, a priori. It is simply to point out that their apriority is owed to nothing deeper than convention or definition.

Kant claims, by contrast, that I find myself bound by rules not of my creating—rules conformity to which come with the nature of thinking (the principle of non-contradiction) and with the very activity of extending my thinking to objects of experience outside me (space and the categories of magnitude as the form of singular reference). Here the element of choice and, so, the question of deviant rules does not arise. Nor need we pause over the suggestion that the activity in question—thinking or experiencing that thus-and-so—can be constituted only by appealing to a richer cultural situatedness. We invite such rejoinders as these when we suggest that (4) defines the activity of thinking[6], or that (5) and (6) define the form of empirical cognition. By placing the emphasis on definition we suggest a degree of control these cases do not allow. Unlike in the Searlean contexts, there is no suggestion that we create the connection between non-contradiction and

[6] Pollok neatly articulates the view against which I am cautioning. He writes that, "It may be true that, for Kant, the laws of pure general logic are as analytic for any kind of understanding as the moral law is analytic for a holy will […]. Accordingly, in these two cases the laws could be seen not as imperative, but rather as defining the activities of thinking and holily willing respectively" (Pollok 2017: 8-9). Here Pollok is imagining someone following Searle's procedure: we incorporate the logical laws into the subject concept—here, "understanding"—so that they are then contained analytically within it, and so are, in this attenuated sense, available a priori. I am tracing (as does Pollok) a more robust, Kantian route to apriority, namely, via the possibility of experience.

thinking; rather, it arises out of the nature of the subject matter at issue. Nor, for the same reason, can we pretend to cement, through definition, the connection between spatializing and experience. Rather, the legitimacy of following such rules as (4), (5), and (6), is a priori in the strict Kantian sense, that is, absolutely independent of all experience.

5. Social Ontology

Talk of rule-following seems to bring with it a context of subjectivity; it invites us to focus on the one following the rules. But it is on just this point that the full weight of Kant's Copernican revolution in epistemology comes to bear. In the *Critique*, Kant puts the matter this way: "The conditions for the *possibility of experience* in general are at the same time conditions on the *possibility of the objects of experience*, and on this account have objective validity" (A 158/B 197, original emphasis). That is, having shown that the rule-following at issue yields knowledge (indeed, a priori knowledge) of objects outside me, enduring through time, undergoing changes, and causally interacting with other such objects—knowledge of such things as chess pieces and stop lights—Kant seems to feel he has put to rest any doubts about the robustness of what he calls his "empirical realism".

As concerns the doubts, Kant could not have been more wrong. Searle speaks for many, in Kant's day and ours, in alleging that Kant perpetuates "the single greatest disaster in the history of philosophy over the last four centuries"—namely, "the rejection of the idea that our perceptual and other cognitive equipment gives us direct access to the real world." Searle continues:

> Kant […] rejected direct realism as a theory of perception. By direct realism I mean the view that, at least on some occasions, in the veridical cases, our perceptual apparatus and other cognitive equipment gives us direct access to objects and states of affairs in the world. We literally see that it is raining or that there is a table in front of me (Searle 2010: 407-408).

As indicated by my discussion thus far, I think this represents a fundamental misunderstanding of Kant's theory of empirical cognition. It never occurs to Kant to doubt, in the veridical cases, that it is raining or that there is a table in front of us. The puzzling thing, for Kant, is how we can have a priori knowledge of such things as rain and tables when that knowledge is not simply a matter of breaking the subject concept down into sub-concepts according to the principle of non-contradiction. In reflecting on this puzzle, we arrive, as above, at the notion of rules conformity to which makes empirical cognition possible. From there we are driven to the idea of objects considered apart from our conformity to these constitutive rules—the idea of *Dinge an sich selbst*. As far as I can see, none of this threatens the basic claim that we are struck (directly) in sensation by raindrops and tables.

Elsewhere, I have argued that we can extend Kant's empirical realism into an area in which Searle has made large contributions, namely, that of social ontology. I have argued that we cognize at least some social groups in the same way as we do chess pieces, stop lights, rain, and tables. Which social groups? Those that reflect a high degree of shared intentionality. In brief: If the bonds between members of a group are such as to legitimize viewing the group as acting and reacting as one, then we may be justified in seeing it as a single dynamical whole.

For example, a soccer team composed of novice players will be seen, upon inspection, as eleven persons acting and reacting independently. But assuming a high level of play, and assuming the deep level of cooperation and interlocking plans and intentions between teammates that make it possible, we may cognize the team as a single, dynamical whole made possible through the cognition of its parts. Here we may think of looking at a Calder mobile—of the moment when we see not just a collection of separate objects, but, again, a single dynamical whole (Godlove, "Perceiving Social Groups: A Kantian Account," under review).

To be sure, taking a Kantian line in social ontology will require limiting referential import to those entities whose cognition reflects both mathematical and dynamical elements. For the same reason that numbers and circles do not exist, neither do marriages, banknotes, academic degrees, debts, rewards, punishments, and recessions. Not that our talk about such matters is somehow suspect; on the contrary, such talk obviously is legitimate, even indispensable. Rather, the burden is on the Kantian social ontologist to make sense of each item on this list (and of course many more) in such a way as to dissolve the appearance of referential import. While he is not writing with Kant in mind, Frank Hindricks adopts, in his discussion of debts, just this kind of deflationary stance. Instead of populating the world with debts we can recognize that what exists are persons whose actions have consequences—sometimes including the owing of another party a sum of money (Hindricks 2008: 137). And so for consequences, owings, and so on. Notwithstanding Searle's dim assessment of Kant's contemporary value, if we are thinking with Kant, we can help ourselves to Searle's work on "status functions" to make sense of such things as money and private property. Paper and bicycles exist; paper becomes money and bicycles acquire owners, as Searle puts it, "in virtue of [their] collective acceptance or recognition" (Searle 2010: 423).

6. Imputability

I have been developing a Kantian line of thought on which certain rules governing thought and experience display a constitutivity so deep as to be uncharacterizable in normative terms. Because they make possible the activity of thinking or experiencing that the world is thus-and-so I cannot both fail to conform to them and retain my identity as a cognizing subject. Since Searle's constitutive rules apply to or (help) make possible discrete domains of cognitive activity—playing chess, driving, etc.—they depend upon the Kantian rules that make cognition possible.

Let us now consider an objection from, as it were, behind Kantian lines. In *Kant's Theory of Normativity,* Pollok has recently argued that the Kantian rules I have canvassed can be constitutive of thought and experience and, at the same time, be accorded normative status. He agrees that the rules in question fail Tolley's test for normativity, but he calls for a "modification" of the criteria. Here is the central passage:

> A modification of Tolley's account of 'constitutive' is required partly because the subject must be taken to be accountable for her performance, such that, even if she articulates a correct sentence, it does not count as a judgment or experience, a moral maxim, or a judgment of taste, unless it is imputable to her […]. Even a toddler with no clue about the rules of chess makes a valid move if she moves the rook horizontally. By contrast, when it comes to judgments what is important is not the *compliance* with principles *simpliciter* but the *imputable compliance*. Unlike moves in chess, judgments do not merely accidentally conform to those principles.

> They are rather liable to assessment in light of them. Merely emulating the words of her parents, even if the result has the form of a correct sentence, does not yet amount to a *judgment* imputable to the child. The sentence is not liable to assessment in light of synthetic a priori principles, because the child is not yet seen as a subject standing under those principles. Thus Kant's synthetic a priori principles may be constitutive *and* normative insofar as our judgments are liable to assessment in light of them. The use of our reason, generally speaking, is normative if in a certain kind of cognitive activity we have a self-understanding of what we are doing that guides the activity […]. This self-understanding commits one to the relevant constitutive principles, which turns them from constitutive to normative, or, more precisely, explains how they can be both (Pollok 2017: 9-10).

Pollok's motivating thought here is that making a judgment requires more than simple compliance with a rule; so much could be accomplished through natural necessity or, as in his example of the toddler making a chess move, happenstance. His candidate for the "more"—developed at length later in the book—is the right kind of self-understanding, one in which reason "guides the activity". Pollok thinks that, emended in this way, Tolley's criteria will show the Kantian rules we have canvassed to be constitutive of and also normative for thought and experience.

As the quoted passage makes clear, Pollok is surveying a landscape considerably wider than the one I have in view. Whereas I am commenting only on Kant's theoretical philosophy, Pollok also has in view Kant's practical philosophy and his aesthetics. No doubt the notion of being led by one's reason or self-understanding plays a central role in these latter two areas. The question is whether, as concerns Kant's epistemology, the situation is fundamentally different.

Our inquiry thus far has been resolutely first-personal. I have been adopting the stance of one who finds that he or she is bound by rules such that a violation of them renders thought and experience impossible. That is the consideration that supports the claim for the kind constitutivity that interests Kant. Note that, in the above passage, Pollok is asking about the development of a child. This switch to the third person marks a departure from the line of thought we have been pursuing. If we agree to purely behavioral criteria, then Pollok is right to say that a toddler can legitimately make a move with a rook. She can accidently, but not imputably, comply with the rules of chess. More generally, he is correct that such criteria cannot license our imputing to the child conformity with those Kantian rules underlying thought and experience canvassed above. We see the child point to a rook and make the sound, "Rook!". If, by this vocalization, we take her to be referring to the rook at issue, then we might—if the child is of a certain age—be tempted to conclude that the Kantian rules are in force. For how, we might reason, could she, at one and the same time, intend and not-intend to refer to the rook? And she must be locating the rook at some distance from where she finds herself, since so much is required by the practice of ostensive reference. But when we picture the alleged impossibilities we are actually running up against nothing having to do with the child. In concluding that he or she cannot both mean and not-mean this or that I have in view my own inability to violate, with impunity, the principle of non-contradiction. In taking her to have no choice but to locate the rook at some distance from where she finds herself I am reflecting nothing more than that inability in my own case. With the third-person stance we may appear to be imagining the child making vocalizations and movements, but the

Kantian constraints in play have to do rather with what I am able to imagine her thinking and what makes singular reference possible for me.

7. A Priori Truth

By adopting a third-person stance, Pollok wishes to ask whether the child's rule-following is or is not guided by his or her reason. This move is tempting, because at some point in the child's development it makes sense to ask whether his or her verbalizations and movements are so guided. And having made this question legitimate we then have the distinction between compliance and imputable compliance, and, moreover, we may well have made room for what I have been denying, namely, Kantian a priori rules that are both constitutive and normative. But since our inquiry is unavoidably first-personal, we must redirect Pollok's question about the toddler. The question becomes: Am I following the rules that make thought and experience possible in such a way that I can impute the rule-following to myself? To find in the positive I must, following Pollok, find that my reason is guiding the activity.

In my own case, however, there is no room for Pollok's distinction between compliance and imputable compliance. I find myself bound by a rule—the principle of non-contradiction—that makes thinking possible. Suppose I ask myself whether my thinking, bound in this way, is governed by my reason. I see at once that I cannot both think that my thoughts are governed by my reason and think that they are not governed by my reason. But then, as above, it follows that Tolley's criteria of normativity cannot be met in this case. I cannot give sense to the idea of violating this rule and remaining a thinking creature. It is an a priori truth in the radical Kantian sense: true absolutely independent of experience because presupposed by it. At this point, the reply might be: We are not doubting that we must comply; the question is whether our compliance is governed by reason. But this requires me to externalize my thinking. That is, when I try to inspect my thinking in the requested way I am trying to treat my thinking as though it were being undertaken by someone else. In other words, it requires that I adopt a third-person perspective on myself, which, as we have seen, would violate the terms of our inquiry, and which, in any case, is not something I can do.

What of the rules governing singular representation—that, for example, I must locate objects at some distance from me? Here I am able to at least think its violation—to think, say, about avoiding checkmate by moving my king through a fourth spatial dimension. And this conceptual latitude may seem to open up the space to inquire whether my compliance is or is not guided by my reason. Here again, it will not do simply to observe that I cannot actually make this move, that, as soon as I reach (physically or, blindfolded, in imagination) for the rook I presuppose its location at some distance from me—thereby requiring of the rook exactly three spatial dimensions. This reply will not do, because, as above, the challenge we are now considering is not whether I am so constrained but whether my compliance with the constraint is imputable. The proper reply is that the fact that I do locate objects of experience at some distance from where I find myself makes it possible for me to inquire into the nature of the rules governing singular reference in my own case—including the requirement that I locate objects of experience at some distance from myself. I am again being asked to investigate the nature of an epistemic constraint as though I were able to view myself as another. But in fact I cannot undertake this or any such investigation without following

the rule I claim to be investigating.[7] All I can say is that either I comply or the possibility of experience collapses.

It results that, in the context we are pursuing, the question of imputable compliance can only be raised by one who asks it of herself—at which point the distinction between compliance and imputable compliance can no longer be maintained.[8] One reaction to this result would be to generalize the problem of imputability; having been unable to vindicate the kind of rule-following we have had in view as guided by my self-understanding I may then come to doubt that my thoughts are meaningful—to become uncertain whether I am merely uttering strings of word-sounding noises rather than imputable judgments. This reaction, a form of radical skepticism, I cannot take up in this paper. Another reaction would be a strict naturalism, perhaps of a Deweyian or Quinean sort. Here I would treat the toddler and my own case as on par, reasoning that what holds for her holds as well for creatures like her. But this would be to pretend that my thoughts and experience are not bound in the ways we have been canvassing, to represent myself as able to philosophize from a point of view that does not already reflect their legislative force. If, on the other hand, we hew to our Kantian line, we will recognize the rules at issue as constitutive of thought and experience even though they are not to be found within their respective general concepts. That is, their connection to thought and experience is neither analytic nor normative, but comes rather by way of the possibility of experience: they must be presupposed if thought and experience are to be possible. It is in this spirit—one of recognition or acceptance—that I think we should understand Kant's cryptic remark that the categories are "self-thought [*selbstgedachte*] first principles *a priori* of our cognition" (B 167).

8. Conclusion

I have developed a train of thought on which a priori rules emerge the following of which is constitutive of, but not normative for, thought and experience. And I have contrasted the resulting Kantian sense of a priori truth as independent of all experience, because presupposed by it, with the use Searle makes of a priori truth by stipulation or definition. There need be no antagonism between these two very different senses of apriority. On the contrary. Searle is presupposing a thinking

[7] Kant was quite clear on this point, noting, in the case of the causal principle, that it "has this peculiar character that it first makes possible its own ground of proof, namely experience, and must itself always be presupposed" (B765). For discussion, see Baum 1979.

[8] I have been arguing that a Kantian stance in epistemology rules out the distinction Pollok wants between compliance and imputable compliance, but my discussion is incomplete in two significant respects. First, Pollok appeals partly to the special case of what Kant calls "transcendental illusion". This occurs when we apply the concept of existence outside its legitimate arena—that is, outside of what strikes us, directly or indirectly, in sensation. Kant claims that this yields a characteristic form of illusion, in which context we are, as Pollok puts it, "liable to assessment," even though the faulty judgments are, in some sense, involuntary (Pollok, *Kant's Theory of Normativity*, 10). Second, Kant defends our reliance on so-called "regulative principles" (A642/B670 ff). Like the rules I have canvassed, these are said to "make experience possible," but they can be given only a practical and not a theoretical warrant. It may be that these principles are both constitutive of and normative for experience. While both of these topics would have to be included in a full treatment of the present subject, I do not think their absence affects the conclusions I draw here. I discuss the second at length in Godlove 2013.

subject, one who must, for example, locate objects of experience—including, as I suggested in section 5, those populating the realm of social ontology—at some distance from him or herself. Thus, Searle is presupposing conformity to the Kantian rules. Nor is there any question of converting the Kantian rules into Searle-style definition-based rules. If I want to think that thus-and-so or cognize some aspect of the world as thus-and-so, I must conform to the higher-order rules canvassed above—and so they will be legislative for any and all of the objects that Searle, or any of us, has in view. In fact, by focusing on the synthetic a priori rules of thought and experience we can make good on the sense of constitutivity that Searle had in mind from the beginning. By virtue of their apriority, the Kantian rules are able to do what the constitutive rules of football and chess cannot: they are able to define the nature of an activity, namely, thinking or cognizing that thus-and-so. They tell us, independent of their cultural or social context, what kind of activity results from our compliance with them.[9]

References

Baum, M. 1979, "Transcendental Proofs in the Critique of Pure Reason", in Bieri, P., Horstmann, R.-P. and Krüger, L., *Transcendental Arguments and Science: Essays in Epistemology*, Dordrecht: Reidel, 3-26.

Bierman, A.K. 1972, "Chessing Around", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 23, 1-2, 141-42.

Cappelen, H. 2011, "Against Assertion", in Brown, J. and Cappelen (eds.), *Assertion: New Philosophical Essays*, Oxford: Oxford University Press, 21-48.

De Jong, W.R. 1995, "Kant's Analytic Judgments and the Traditional Theory of Concepts", *Journal of the History of Philosophy*, 33/4, 613-41.

Godlove, T. 2013, "The Objectivity of Regulative Principles in Kant's Appendix to the Dialectic", in Ruffing, M., La Rocca, C., Ferrarin, A., Bacin, S. (eds.), *Kant und die Philosophie in weltbürgerlicher Absicht: Akten des XI. Kant-Kongresses 2010*, New York: DeGruyter, 129-40.

Godlove, T. "Perceiving Social Groups: A Kantian Account", under review.

Hindricks, F. 2008, "The Status Account of Corporate Agents", in Schmid, H.B., Schulte-Ostermann, K., Psarros, N. (eds.), *Concepts of Sharedness: Essays on Collective Intentionality*, Piscataway: Ontos Verlag, 119-44.

Kant, I. 1998, *Critique of Pure Reason*, New York: Cambridge University Press.

Kant, I. 2004, *Lectures on Logic*, New York: Cambridge University Press.

Lorini, G. 2012, "Can Constitutive Rules Create a Practice?", *Praxis Filosófica*, 34, 5, 139-48.

Lorini, G. 2014, "Meta-institutional Concepts: A New Category for Social Ontology", *Rivista di Estetica*, 54, 127-39.

Lu-Adler, H. 2017, "Kant and the Normativity of Logic", *European Journal of Philosophy*, 25, 2, 207-30.

[9] For putting the point in this way I am indebted to Bierman 1972: 141. For helpful comments on an earlier version of this paper, my thanks to Tony Dardis, Konstantin Pollok and to two anonymous reviewers for this journal. Any remaining errors are of course my own.

Melnick, A. 1985, "Kant's Theory of Space as a Form of Intuition", in Kennington, R. (ed.), *The Philosophy of Immanuel Kant*, Washington D.C.: Catholic University Press, 39-56.

Mosser, K. 2008, *Necessity and Possibility: The Logical Strategy of Kant's Critique of Pure Reason*, Washington D.C.: CUA Press.

Pollok, K. 2017, *Kant's Theory of Normativity: Exploring the Space of Reason*, New York: Cambridge University Press.

Searle, J.R. 1969, *Speech Acts: An Essay in the Philosophy of Language*, New York: Cambridge University Press.

Searle, J.R. 2013, "Reconciling the Basic Reality and the Human Reality: Post Kantian Themes", in Ruffing, M., La Rocca, C., Ferrarin, A., Bacin, S. (eds.), *Kant und die Philosophie in weltbürgerlicher Absicht: Akten des XI. Kant-Kongresses 2010*, New York: DeGruyter.

Schwyzer, H. 1969, "Rules and Practices", *The Philosophical Review*, 78, 4, 451-67.

Shenk, D. 2007, *The Immortal Game: A History of Chess*, New York: Anchor Books.

Thompson, M. 1981, "On A Priori Truth", *Journal of Philosophy*, 78, 8, 458-82.

Tolley, C. 2006, "Kant on the Nature of Logical Laws", *Philosophical Topics*, 34, 1-2, 371-407.

What Does it Mean that Constitutive Rules Are in Force?

Bartosz Kaluziński

Adam Mickiewicz University in Poznań

Abstract

The aim of this paper is to shed some light on the issue of how we can understand constitutive rules as being in force for participants *S* in some rule-constituted practice. We take a look on complicated team-games that are broadly conceived as model example of rule-constituted practices. We claim that rules of games are dependent on mental states of participants in that practice. More precisely, they are in force for the participants *S* of such games if these participants jointly meet the following conditions: (1) every *S* has a working knowledge of the rules, (2) every *S* intends to and actually conforms to the games' hard core rules and (3) every *S* respects the remaining rules (i.e. in the case of an alleged rule violation a player asks him or herself what decision an ideal referee would have made in those circumstances, and conforms to that decision).

Keywords: Constitutive Rule, System of Rules, Existence of Rules, Games, Ideal Referee.

Introduction

John R. Searle (1969: 33-41) explicitly made the famous distinction between 'regulative' and 'constitutive' rules (it should be noted that Searle was neither the first nor the only philosopher to use a distinction of such type, as von Wright (1963) or Rawls (1955) also stressed the difference between these rules, yet the Searlean account seems to be the most influential one). The notion of constitutive rules is widely discussed in the philosophy of language (e.g. Gluër and Pagin 1999, Williamson 1996) or in the philosophy of law (e.g. MacCormick and Weinberger 1986, Marmor 2009). Roughly speaking, regulative rules are imperatives, i.e. they tell us what to do or how to behave in certain circumstances. A paradigm example of regulative rules are the rules of etiquette, e.g. "During an official dinner, officers should wear a tie". Rules of this kind regulate antecedently existing activities or practices. This means that it is possible to describe these practices independently of the existence of such rules. We can describe a dinner as a dinner even if there are no rules of etiquette concerning what one should wear during that dinner. Rules of this kind will not be a part of my investigation in this paper.

Argumenta 4,1 (2018): 111-124
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.kal

Now let us have a look at Searle's constitutive rules. These rules are not imperatives (according to Searle they do not have an action-guiding function) but instead 'create' new forms of activities or practices.[1] Constitutive rules are metaphysically prior to the practices or activities they 'create', i.e. some practice or activity could not exist at all without these rules. How is this to be understood? Suppose that twenty-two people are running around on the grass and making such moves as the members of football[2] teams usually make. Is it possible to describe their behaviour as that of a football match? The answer is in the negative, unless we know that the football rules are *in force* for these people. Playing football is not a matter of, to use Searle's words, the 'brute facts'—that somebody is kicking a round object or hitting it with his/her head. The existence of football rules makes it possible to describe the moves of twenty-two people as playing football.

Of course, it is not the case that constitutive rules are simply 'out there', that we can take their being in force for granted without caring about some social or psychological context that is necessary for these rules to appear. As noted by Pagin, "[t]he being in force of a rule depends ultimately on acceptance, on regarding the rule as in force" (1987: 15), because rules are not physical objects, they do not have their very existence regardless of any recognition by the subjects of these rules. In my paper I will not investigate closely that deep metaphysical question of recognition of constitutive rules, I am not particularly interested in the question of what brings rules into existence in the 'last instance'. Rather I will focus on the more prosaic issue of what is the relationship between any given Jane's or Joe's (who is a participant of a rule-constituted practice) commitments (motivations, intentions etc.) towards previously enacted rules. I take the existence of rules for granted and then ask what needs to be true for any given Jane or Joe in order to justifiably claim that she or he is a participant of a practice constituted by rules. At first glance, there are three main possibilities for the relationship between a set of constitutive rules *R* and a participant *S* in some practice *P* (cf. Pagin 1987: 12-33; Gluër and Wikforss 2015):

1. Participant *S* is motivated (intends, is committed) to act in accordance with *R*,
2. There is no link between *R* being in force and *S*'s mental state (*R* is in force regardless of *S*'s intentions, motivations, commitments),
3. *S* is motivated (intends, is committed) to act in accordance with only some subset of the set of rules *R*.

At first glance, it might appear that some form of 2. is the most plausible option, but is it the case? In order to answer this question I will look carefully at some

[1] As Gluër and Wikforss claimed, "[a]ccording to Midgley (1959) and Searle (1963, 33ff), constitutive rules typically, and naturally, can be put into the following form:

(CR) In C, doing X counts as doing Y.

[...] [T]his characterization of constitutive norms is too narrow – there are prescriptions that are constitutive of certain games (for instance, it's constitutive of ice hockey that spearing is forbidden). A wider characterization of constitutive norms or rules thus counts rules or norms as constitutive of a certain action, or activity, *A* iff *A* cannot be performed, or engaged in, unless these norms are in force" (2015).

[2] The term for this game is 'football' as used and understood in Europe or South America, not in the USA, i.e. to denote association football.

complicated practices established and governed by rules. The paradigm example of such practices are games. Philosophers have hitherto mainly analysed the game of chess (e.g. Marmor 2006), which seems to me not to be the best choice. The game of chess comprises a relatively small number of rules and admits no violation of its rules (e.g. moving a rook diagonally seems to be something "outside the game" or even entails the termination of the game). Hence, if we take a game of chess as our model we will come to the conclusion that a participant in the game of chess incurs commitment to act in accordance with all (or almost all) rules of that game. Nonetheless, we have rather a strong intuition that the rules of games are quite frequently violated (there are multiple fouls and other 'illegal' moves within the game of basketball or football), so the game of chess, which is unforgiving when it comes to violating its constitutive rules might not be representative for rule-constituted practices at all. I am convinced that if we look carefully at more complicated games (e.g. basketball, football, rugby) that easily allow for a violation of rules without termination of the game, it is possible to obtain more credible knowledge about the relation between participants (their intentions, commitments etc.) of rule-constituted practices and constitutive rules, hence we will know more about the social and psychological context within constitutive rules operate.

1. Why Focus on Rules of Games instead of Legal Norms?

One may ask if constitutive rules in general are very different from, for instance, legal norms. It is not difficult to notice that legal norms also perform the constitutive function, i.e. without some set of legal norms there would be no such institution as a limited liability company, pledge, mortgage or inheritance. Sometimes legal norms actually create some forms of activity or practice, e.g. inheritance, which is the practice of passing on property, debts and obligations upon the death of an individual. We have a system of precise rules that determine who is entitled to receive a share of the deceased person's property or obligations and under what conditions, yet obviously there are serious differences between legal norms constituting such an institution as inheritance and constitutive rules in general. What are they?

First, legal norms are in force for all people regardless of their attitudes, desires or intentions. When one dies, his or her property, obligations and debts are passed on to some other subjects due to inheritance law. Of course, the norms of inheritance law may, and in fact usually do, make it possible to write down a will. We are normally allowed by the authorities to express our will concerning the distribution of our property after our death. However, one must note a critical fact—we can create our will only because inheritance law, enacted in a proper procedure by the proper authorities, allows us to do so.[3] Hence, when it comes to such a legal institution as inheritance we only have a certain amount of liberty that was given to us by the legal norms (e.g. in some countries, such as France, there is a system of forced heirship that allows the testator to freely dispose only half of his or her property; in many other countries the testator is not subject to

[3] Of course, it may be the case that something similar to the practice of last will preceded any legal systems, as a part of a custom or some religious system. But it was not the same institution as the contemporary testament, which is, obviously, determined by legal rules.

such a limitation). Of course, we can change the legal system we are subject to by simply moving from one country to another, but we are always subject to some legal system (when we are on international waters or in space the Law of the Sea or Space Law is respectively in force). Our mental states play no role in this situation—these norms are in force whether we like it or not, but in the case of other practices or institutions created by rules, such as chess, our intentions or desires do indeed play some important role, i.e. we can choose whether we would like to participate, for instance, in a game of rugby, but we cannot choose to 'participate' in a legal system.

Second, legal norms are *usually* associated with the whole apparatus of coercion. In the case of the violation of a legal norm there is a wide range of instruments that the state can use to punish the offenders. The police can use force; we have a criminal court trial that can convict the offenders and a penitentiary system that can imprison them—all of these could be legitimately applied by the state to bring order and the rule of law. In the case of other 'institutions' constituted by rules, things seem to be different. For instance, when we violate a rule of a game, we will not be fined, prosecuted or put into jail.

Legal norms can sometimes perform a creative function, but they are not a good model for constitutive rules in general, because they have some features that seem to be unique to them, for instance, the whole apparatus of coercion legitimately tries to ensure that the law is not violated. Moreover, there are intimate connections between law and politics and law and morality. All these factors seem to be absent in the case of games (and probably in the case of other rule-constituted practices, e.g. types of illocutionary acts). Hence, it appears that despite there being broad literature concerning the constitutive character of legal norms, these considerations cannot be easily used to make claims about constitutive rules in general. Therefore, it appears that our plan to investigate the rules of games in order to acquire knowledge concerning constitutive rules in general seems to be justified. Analysis of rules of games, that are broadly conceived as great examples of rule-constituted practices, can give us a portion of knowledge on constitutive rules that might be subsequently extrapolated to some issues in social ontology or philosophy of language.

Let me state clearly at this point, I do agree with the claim that many legal norms are constitutive. I only have doubts as to whether these rules are representative for constitutive rules in general, and so whether any knowledge that comes from analysis of legal norms might have very limited application to other rule-constituted practices.

2. Constitutive Rules and Commitments

We have noted that there are some differences between constitutive rules in general and legal norms. The former are, in some sense, dependent on one's mental states, and the latter are independent of the subject's mental state[4] (one cannot say

[4] Of course, in some sense constitutive rules are *ultimately* mind-dependent. For instance, the rules of football that are listed in official rulebook would not come to exist if there were no will to enact them by certain people that work in FIFA. Or to phrase it differently, there would be no rules of football, or football matches in their present form without the decision to formalise and enact the game's constitutive rules, by relevant associations. Let me

"I do not have intention/desire/motivation to follow the rules of highway code, so they are not in force for me"). It seems that, despite being somehow dependent on one's mental states, constitutive rules are not to be changed freely or invalidated by one's intentions or desires. Now is the time to look closely at games such as football or basketball in order to figure out what the relation between one's mental states or commitments and constitutive rules looks like.

The first question I would like to address is whether participants in a practice are committed to conform to its constitutive rules. One may think that the very concept of participation in some activity without being committed to the rules of that activity is bizarre. It seems natural to think that when we play, for instance, football, we are committed to act in accordance with its rules; but this is not always the case. Consider the notion of a 'tactical foul', which is an intentional violation of the rules of the game in order to interrupt an opponent's action (that kind of rule violation can be considered an "illegal" move within the game, that is it is a violation that brings consequences within the game—penalties, free kicks, yellow cards; the difference between a tactical foul and a usual one is that the former is made deliberately and its purpose is to prevent the opposite team from scoring a goal whereas the latter is usually a result of lack of skill). Sometimes it is beneficial to foul an opponent by, for instance, holding him or her by the shirt and thus preventing the opponent's team from scoring points (in the case of basketball) or a goal (in the case of football). Hence it is clear that playing 'correctly' (i.e. in accordance with all the rules of the game) is sometimes less effective than playing 'incorrectly' (i.e. by violating some rules of the game in order to win). There are many rules that could be violated, one after another, without termination of the game. Suppose that a basketball team's players perform the following consecutive actions: travelling, double dribble, lane violation, five-second violation, and carrying. During these five consecutive actions they have violated five different rules of basketball, but it is not the case that with every such violation they began to stop playing basketball and started to do something else. They are just terribly bad players.

It seems that some player *S* can play a game and violate the rules of that game either via some sort of inability or intentionally. The violation of rules that occurs because of the players' mistakes or lack of skills is quite uncontroversial; it is reasonable to think that when both in the case of football and basketball the defending player is trying to pick up the ball from the player of the opposite team but fails to do so because he or she is, for instance, not fast enough, that defender actually does not question the rules that forbid tripping or pushing his or her opponent. There is no huge mystery here. The case of the so-called tactical foul is far more interesting because it proves that it is *not necessary* for a player to be committed to or to intend to conform to all of the rules of the game; on the contrary, the player has the intention to violate the rule(s).

At this stage the following question arises: Can a player intentionally violate any rule of the game? I believe that the answer is negative. It seems reasonable to assume that the constitutive rules could differ from one another and the system of

remind you that in this article I am not particularly interested in above-mentioned origin of the rules or ultimate 'rule of recognition', rather I am investigating the relation between states of mind of participants (commitments, intentions etc.) and the rules (after they have been enacted).

constitutive rules of the game consists of rules, so to speak, of different importance. There is a 'hard core' of the game, which is a very small number of rules that cannot be violated and a much larger set of rules that could be violated without any *serious* consequences. Violation of the rules of the former type, i.e. those rules that form the hard core of the game, entails termination of the game. Violation of the rules of the latter type is, therefore, considered a fault, a somehow 'incorrect move', but this would not entail the fact that we terminate playing the game. An example of such a constitutive rule of football could be the rule that forbids holding one's opponent by the shirt. Behaviour that counts as violating such a rule is considered to be a foul, but such an act certainly does not mean that the player terminates playing football. There is no risk of termination of one's participation in the game, but there are other, far less serious, consequences—free kicks, yellow cards etc. This is because rules are interconnected—a rule that forbids kicking the opponent is interconnected with the rules that states what the consequences of kicking the opponent are.

However, things are different in the case of rules that form the hard core of the game. It appears to me that it is essential in such a game as football that there be a goal and that we be trying to put the ball into it by using our legs or head; in the case of basketball we are trying to put the ball into the basket by using our hands. If some football player started passing the ball to his or her teammates, and *not* with his or her foot but with his/her hands, he or she would, of course, terminate playing football, i.e. it would just not be a football match anymore. If someone started kicking the ball into the basket this would not be a basketball game anymore either. Now it is clearer what it means to play a certain game—we play a game if we act in accordance with the rules that form that game's hard core. If we violate some of its other rules we just make a foul, make an 'illegal move'. Hence players are committed to only a very small number of rules that constitute the game, and they sometimes have the intention to violate the rest of its rules.

3. The Role of Deep Conventions

But is it the case that a mere lack of violating a small set of the most important rules is a sufficient condition of playing a game? In my opinion the answer is in the negative. It is possible to play a certain game without intending to act in accordance with *all* of its rules, but surely it is not possible to play a game without intending *to play it*. What I mean by this is, basically, the same as the remarks that were expressed by Hubert Schwyzer:

> Promising and asserting and the others are *practices* or *institutions*, like chess or basketball, or making a will, or performing civil matrimony. They are *systems of rules*, and the rules *define* them. Alston, Rawls and Searle make much of this analogy with games and institutions in their discussion of the concepts of these sorts of acts. [...] I shall argue that in the relevant sense rules *do not*, and *cannot*, *define the nature of an activity*; the rules of chess, for example, *do not explicate what it is to play chess* (Schwyzer 1969: 452-53).

If I get things correctly the argument that Schwyzer provides on the pages following that passage is as follows: rules of the game cannot define the nature of the activity which is constituted by those rules. In the case of games it is possible to

perform all of the moves or actions that the players typically make without playing the game in question. It is not so difficult to imagine that a group of people is behaving like players, but after the game is over one 'team' looks relieved and one is terrified because of the 'result'—they believe the gods will punish them or something of the sort. The crucial point is that the mere fact that some people behave like players normally do, i.e. that their actions are in accordance with the rules of the game, is insufficient to see them as players because they might perceive their activity as a part of some sort of, for example, religious ritual. People who do not have the concept of a game, particularly of sport, cannot participate in the game of football, basketball or any other. A very similar point was attributed to Michael Dummett (1978) by Anandi Hattiangadi:

> Dummett draws an analogy between the concept of truth and that of winning at a game such as chess. He observes that you could specify all the rules of chess, specify how all the pieces move, what constitutes winning or losing the game, and still something would be left out: that the whole point of the game is to win (Hattiangadi 2009: 191-92).

Generally speaking, I agree with the remarks made by Schwyzer and those attributed to Dummett. Constitutive rules alone are not sufficient to establish any 'institution', there must be something in the social background that, together with the rules, makes the 'institution' possible to appear—for instance it is not possible to play football without having the concept of playing a competitive game (cf. Marmor 2007). Below I intend to provide a few examples that I hope might make these points slightly clearer.

Suppose that you are a spectator of the boxing fight between Muhammad Ali and George Foreman that took place on 29 October 1974 in Kinshasa, Zaire. There are about sixty thousand people watching the fight in the stadium. All of these people are watching Ali and Foreman punching each other. Now you can ask yourself why nobody is calling the police.

Now suppose that you are is in the stadium watching a football match. The player of the first team makes a tackle in order to take the ball from his or her opponent and he or she fails to do so; moreover, the consequences of this very poor tackle are huge—the player of the opposite team gets seriously injured, e.g. his/her Achilles tendon is torn and he or she needs to undergo surgery and some rehabilitation. Why is the player that made the unfortunate tackle not subject to criminal charges concerning serious damage to his/her opponent's body? Why does the injured player have no possibility to press charges against the unfortunate defender for reimbursement of the rehabilitation costs?

Or to provide you with another example of the importance of *deep conventions*, suppose that basketball players actually do not commit any fouls, there are no lane violations etc, but they do not make any attempt to score points, their moves do not violate any rules of the game, so that the referee has no justification for any intervention but still they are not participating in the practice of playing the game, because the whole point of the practice of playing a competitive game is to win.

I believe that the above-mentioned examples could help one to understand what the point is. One may know all of the rules of boxing, chess or football, and yet one will not understand many crucial things, e.g. that the point of the game is to win. Moreover, the game is something different than ordinary life (and of

course it is also distinct from, for example, religious rituals or an art performance) and it has some specific nature. That is why the legal consequences of injuries are different within the game of football or a boxing match and in ordinary life. Roughly speaking, there is something that might be called a *deep convention* that underlies the games (but also other practices—religious rituals, art performances etc). It is a necessary condition of playing the game of football or basketball that the very concept of *playing a competitive game* be known by the members of a certain society. And that deep convention consists in knowing that games are rule-governed activities, the point of the game is winning and that games are quite detached from other types of practices (like art) and ordinary life (cf. Marmor 2006, 2007).

It is now a good moment for a small recapitulation. First, the rules that constitute such complicated games as basketball or football are not very similar to legal norms because (1) there is no apparatus of coercion that ensures the enforcement of these rules or any direct link between rules and morality, and (2) legal norms are in force for all people in a certain territory regardless of one's mental states or commitments while the rules of the games are in some sense mind-dependent. Second, constitutive rules create an activity, they define some actions and determine which moves are allowed, forbidden and required but do not determine the socio-cultural sense that a society ascribes to that activity. It is a necessary condition of playing a competitive game that the very concept of playing a competitive game be known by the members of the society. Third, the participants of a game do not have necessarily an intention or desire to conform to all the rules of that game; however, they necessarily have the following intentions: (1) to participate in the game (i.e. to treat a game as a game, not as a religious ritual, etc.) and (2) not to violate the rules that form the game's hard core.

4. Respect for Rules and the Role of the Referee

Now we know some facts about rules constitutive for games but we mostly know in what respect they differ from legal norms, yet still we know very little about them being in force for the players or, to formulate it another way, what it means that the players are subject to constitutive rules. It is clear that the rules of popular games (football, basketball, rugby, handball, etc.) are listed in the official documents enacted by the respective sports federations. But the very fact that a certain federation issued a document concerning the rules of the game is insufficient to make an assumption that these rules are in force. Paper endures all, but although something is written that does not mean that anybody cares about it. As was noted above, there is the police force or a similar apparatus of coercion that could enforce the legal norms, wherein in football or basketball there is no police that could legitimately use force. There also seems to be no link between morality and, for instance, a set of football rules that define an offside. And, as it was noted above, there is no necessary relation between players intending or desiring to conform to *all* rules of such games such as basketball or football. Hence the question arises as to how it is possible that the rules of the games are in force for the players.

It is not so difficult to notice that in any professional match there is a referee who carefully looks for fouls, awards free kicks, starts and finishes the match, etc. However, he or she does not have similar means as the police and the justice system do. What might be equally important is that when amateurs play basketball or football then, typically, there is no referee. These issues raise the question

of the importance of the referee in the process of rule enforcement. There are two possibilities—either these activities (professional and amateur matches[5]) are not basically the same thing, i.e. a professional and an amateur match only looks similar, but there is some deep difference between them, or the physical presence of the referee is not as important as it looks. Some philosophers like to think that indeed there is an important difference between professional competitions and amateur versions of these games (Marmor 2009). This could be tempting, but I shall argue that in fact the difference between professional and amateur versions of games such as football or basketball are rather insignificant, particularly with respect to the position and physical presence of the referee and the existence of constitutive rules.

First, I would like to say that indeed the figure of the referee is crucial to understanding the nature of the rules that constitute games, but I do not think that we should focus on some individual that is actually on the pitch. All real referees are only human beings, i.e. they become distracted, do not follow the actions, have problems with visibility and make many mistakes. Importantly, they only have a whistle and they do not have guns or truncheons so, as was stated before, their ability to enforce behaviour that is in accordance with the rules is very limited. Moreover, it would be peculiar to assume that their on-pitch decisions in some sense bring rules into being (or to state that there are no rules, just real referees' decisions). Acceptance of the claim that because of the actual referees' decisions the players are subject to the rules of the game would, at best, mean that the existence of constitutive rules is basically non-distinguishable from the existence of legal norms in the sense of the old-fashioned doctrine known as American legal realism (cf. Llewellyn 1962) that reduces the law to decisions made by judges and authorities (and it looks like this doctrine has its best years behind it because of criticism made, among other authors, by Hart 1961 and Dworkin 1977). In the case of rules of games, this would mean that basically all rules are completely irrelevant because the referees' decisions constitute the games. Such a view would entail that all amateur matches without a referee are not matches at all, and that seems very odd.

Hence, it seems that some Joe or Jane who is actually refereeing a match is not the 'source' of the force of constitutive rules—it must be something else. Let us assume that a critical role in bringing the constitutive rules into existence is not played by a physical referee but rather by some ideal referee that is in cognitively optimal conditions, is free of emotions and does not get tired or distracted. This may look bizarre. It was stated that a real person who is refereeing a match is insufficient to ensure rule enforcement and to provide the basis for the statement that the players are subject to rules, so how could such an abstract construct as an ideal referee in optimal cognitive conditions do the trick? This is actually quite simple, as the rules are in force for some player because of his or her *respect* for the rules. But what does this mean? I should like to propose that the crucial point is that it is the expectations of the participants that in some sense constitute the ideal

[5] By "professional" matches I mean those that are part of a system of regular competition under the auspices of a relevant sport association (e.g. Deutscher Fußball-Bund, National Hockey League, Association of Tennis Professionals etc.). The players' participation in such events is what they do for a living, while "amateur" matches are organised by just for fun and health benefits.

referee, although the participants do not need to have intentions or desires to conform to all of the rules and in fact violate these rules quite frequently.

Suppose that you are a part of an amateur football match and you have just grabbed a member of the opposite team by his or her shirt. This was a tactical foul—you did it because you wanted to stop a dangerous action coming from the opposite team. It is clear that you did not have the intention to act in accordance with the football rule that forbids holding opponents by the shirt, but it is still possible to ascribe to you respecting that rule. How is this possible, you may ask.

First, let me state that the necessary condition for respecting a rule is knowledge of that rule. Of course, this is not necessarily an explicit type of knowledge, though children, amateurs or even many professionals actually play, for instance, football without the ability to recite the exact form of all the rules that are listed in the official rulebook. Sometimes we know the exact form of some rules, but in the vast majority of cases we rather have a practical, working knowledge of the rules. I am not going to investigate the subject of knowledge of rules in a more in-depth fashion, I suppose that it is intuitively clear enough for our purposes at this time. Let me go back to the main theme. You held your opponent by his or her shirt in order to prevent the opposite team from making a very dangerous move that could result in a goal. Obviously, the players of the opposite team start to protest—they wave their hands, shout 'Foul!' and so on. The next step is, typically, your team's accepting their protests, i.e. your team acknowledges that your behaviour was a violation of the rule and that the other team should, in consequence, start the game with a free kick or, if your foul took place in the penalty area, with a penalty kick. You and your teammates are willing to acknowledge what you did and to face the consequences.

There is no real referee on the pitch, yet you act as if there were, i.e. the opposite team is awarded a free or penalty kick. You can ask yourself what a fair, unemotional, competent referee would do in a certain situation and then you act in accordance with this 'decision' made by such an 'ideal referee'. This is a similar situation to a hypothetical scenario where there are no road police and people do not risk going to jail or face other sanctions for violating the highway code but still voluntarily pay to some appointed recipient some amount of money each time they drive too fast. In real life this scenario would rather be unlikely. It is not metaphysically impossible but it is not so easy to imagine a society that would work this way. Yet when we play such a complicated rule-constituted game such as football (or basketball), this seems to be the case. To sum up, we do not need to have the intention or the desire to conform to all of the rules of the game. Indeed, we violate them quite frequently, but when such a violation is observed we ask ourselves what an ideal referee would do and act in accordance with his or her decision.

It looks as if we are ready to put things into a clearer perspective. It appears that some player *S* is subject to the set of constitutive rules *R* of some complicated team game *G* if:

1) *S* knows that he or she is participating in the game (knows the deep convention of playing a competitive game) and intends to play,
2) *S* has working knowledge of the rules of *G*,
3) *S* intends to act in accordance with some small subset of *R* that forms *G*'s hard core,

4) *S* respects the other rules of the game, i.e. in case *C* of a supposed violation of any rule that is not a part of *G*'s hard core:
   i. *S* asks him or herself what decision *D* an emotion-free referee in optimal cognitive conditions would have made in *C* (and answers this question),
   ii. *S* conforms to *D*.

It appears that such an account could be viewed as a unification of both amateur and professional types of games—in my opinion they are pretty much the same with respect to the rules and mental states of the players towards these rules. What is probably more important is that it offers an explanation of the statement that players are subject to constitutive rules of games, which are a paradigm example of rule-constituted practices.

5. Reply to Possible Objections

One may have some questions about the account sketched out above; I will try to address some of them. First, one may ask about the relation between the actual referee's decisions and the ones made by the ideal referee. This is, I suppose, quite easy. Decisions made by an ideal and actual referee could diverge because, by definition, the actual referee is not an emotion-free[6] person in cognitively optimal conditions. Hence any real referee can, and sometimes actually does, make mistakes, as only an ideal referee can apply the rules of the game correctly in every situation. This raises the question about the very figure of the ideal referee. One might wonder, given that decisions made by the actual and ideal referee diverge every now and then, if my account is not too 'metaphysical' in that it promises to give an accurate and empirically adequate explanation regarding the question about players being subject to constitutive rules and relies on the very unempirical notion of an ideal referee. Does that kind of reliance not make such an account implausible? I do not think so.

Although the actual person that is refereeing the match has little to do with the ideal referee, that factor seems to be irrelevant. Of course, an actual referee makes mistakes, almost exclusively because he or she is not well positioned, things are happening too quickly or something is impossible to notice with the naked eye. That is the reality, but all of these factors could possibly be eliminated by introducing some electronic devices or making video replays (including slow motion) that would be available to the referee. Proper training and the use of technology (microchips that allow to state if there actually was a goal in the football match or video verification in rugby are already in use) will make the difference between real and ideal referees almost obsolete. Technological progress in the near future will make referees nearly faultless. Moreover, I believe that such an idealisation is permissible in this context. Making idealisations and counterfactuals are just philosophical tools that are commonly used. If one thinks that the figure of an ideal referee is somehow faulty he or she should make an argument against it.

[6] I mean by this that the referee is not carried away by his or her emotions or sympathies, and does not get influenced by the pressure of the audience, media etc. In some cases the ability to read the players' emotions and intentions might be useful for a referee, for instance in distinguishing a deliberately brutal foul from an unintentional one.

Of course, the introduction of technological devices probably would not eliminate all controversies. Some things are hard to judge even when technological assistance is provided, for instance, in order to determine whether a penalty kick should be awarded, we sometimes must decide if the relevant member of the defending team was pushing or merely touching his or her opponent. These situations might be tricky and some disagreement among well trained professional referees that have access to technological devices can occur. But this appears to me rather a question of the vagueness of the terms that occur in the rules. It seems that everybody agrees that, for instance, pushing the opponent is forbidden and in cases when such an event occurs the free/penalty kick ought to be awarded. The disagreement is about whether in the certain situation there was really a push or merely a touch.

Another issue is that players sometimes question the actual referee's decisions or, in the case of an amateur match, the two teams disagree over some decision to be made on the pitch. In order to adequately address that issue one must note that such a situation can have multiple causes: (a) the players do not know a rule, (b) the players question the referee's decisions for tactical reasons, e.g. to put pressure on him or her and, in effect, to obtain a favourable decision, or (c) they do not have proper knowledge of the facts because they are not in cognitively optimal conditions. Let me address these issues.

There are two possibilities if the player does not know the rules, i.e. either he or she does not know some rule at all or, alternatively, does not know the rule sufficiently. In the case of any rule that forms the game's hard core a player that does not know that rule does not play the game. That seems obvious. But sometimes players do not know the rules (the ones that do not form the 'hard core' of the game) precisely enough and only have some working knowledge of the rules, yet they lack the ability to, for example, quote the appropriate passage from the rulebook. This kind of situation might appear every now and then, but it does not seem to undermine the account I have sketched, for two reasons. First, it was stated that players do not need to know the exact form of all rules but rather need to have a practical, working knowledge of the rules. Intuitively, that kind of knowledge leaves room for some inaccuracy, as a working knowledge might just be slightly slack. Second, a player is subject to the rule if he or she respects the rule (i.e. if he/she asks him or herself what decision would be made by an ideal referee). Hence, in this case, he or she does not challenge the rule itself, but rather has a false image of the decision of the ideal referee because of his or her inaccurate knowledge of the rule.

When we look at case (b), i.e. that players are arguing with the referee (or with members of the opposite team in the case of an amateur match) and hoping that the pressure will result in a beneficial decision, it is clear that there is no threat for my account. In this case the players do not actually question the rule itself but are trying to convince others that the facts were different than they appeared, e.g. they are not questioning the rule that forbids kicking the opponent, as indeed they acknowledge that if there was a kick then there should be a free or penalty kick, but they are just pretending that there was no kicking of the member of the opposite team. They are trying to fool the referee or other team in order to win the game.

The last case, (c), seems to be even more obvious than the previous ones. Clearly there is no disagreement regarding the rules, and everybody respects them, but people could have different opinions as to what the decision made by

the ideal referee would be because of their poor positioning, lack of a clear view of the incident, very dynamic character of the event, etc. Hence the dispute is about the facts (for instance, if the member of the attacking team was offside or not) but not about the rule (in this case the set of rules that define the offside). If I am correct, the account of constitutive rules being in force as presented above is able to accommodate common situations that occur during matches, i.e. it seems to be empirically adequate.

6. Conclusions

In this paper I have tried to determine what it means that rules constitutive for certain type of practice are in force for participants of that practice. It appears that the rules of complicated team games (which are paradigm examples of rule-constituted practices) such as football or basketball are dependent on the mental states of the people that participate in these activities, yet they cannot be changed freely. I have argued that the rules of games are in force for the participants *S* of such games if these participants jointly meet the following conditions: (1) every *S* has a working knowledge of the rules, (2) every *S* intends to and actually conforms to the games' hard core rules and (3) every *S* respects the remaining rules (i.e. in the case of an alleged rule violation a player asks him or herself what decision an ideal referee would have made in those circumstances, and conforms to that decision). In sum, once constitutive rules are enacted, some facts about the participants in the rule-constituted practice, as well as about the society in which the practice takes place, must be true in order for these rules to be in force. Conditions (1), (2) and (3), however, cannot arise without a specific social context, that is to say, they are unthinkable in a society that does not know the deep convention of playing competitive games.

Of course, my account of some person *S* being subject to a set of constitutive rules *R* (or, to put it differently, a set of constitutive rules *R* being in force for person *S*) is based on games. They, without a doubt, are good examples of activities constituted by rules. The account presented in this paper relies on the notion of an 'ideal referee'. Though, there are also rule-constituted practices where, at first glance, there seems to be no referee (for instance, types of illocutionary acts), which raises the question as to how precisely an account presented in this paper could be applied to other rule-constituted activities. However, this issue lies beyond the scope of this paper.[7]

References

Dummett, M., 1978, *Truth and Other Enigmas*, Cambridge, MA: Harvard University Press.

Dworkin, R. 1977, *Taking Rights Seriously*, Cambridge, MA: Harvard University Press.

[7] The work on this paper was funded by the National Science Centre, Poland, grant under award number 2015/19/D/HS1/00968. I would like to thank the anonymous referees for comments on the earlier version of this paper.

Gluër, K. and Pagin, P. 1999, "Rules of Meaning and Practical Reasoning", *Synthese*, 117, 207-27.

Gluër, K. and Wikforss, A. 2015, "The Normativity of Meaning and Content", in E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Summer 2015 Edition.

Hart, H.L.A. 1961, *The Concept of Law*, Oxford: Oxford University Press.

Hattiangadi, A. 2007, *Oughts and Thoughts. Rule-Following and the Normativity of Content*, Oxford: Clarendon Press.

Llewellyn, K., *Jurisprudence: Realism in Theory and Practice*, Chicago: The University of Chicago Press, 1962.

MacCormick, N. and Weinberger, O. 1986, *An Institutional Theory of Law*, Reidel: Dordrecht.

Marmor, A. 2006, "How Law Is Like Chess", *Legal Theory*, 12, 347-71.

Marmor, A. 2007, "Deep Conventions", *Philosophy and Phenomenological Research*, 74 (3), 586-610.

Marmor, A. 2009, *Social Conventions: From Language to Law*, Princeton: Princeton University Press.

Midgley, G.C. 1959, "Linguistic Rules", *Proceedings of the Aristotelian Society*, LIX, 271-90.

Pagin, P. 1987, *Ideas for a Theory of Rules*, doctoral dissertation, Stockholm University, https://stockholmuniversity.app.box.com/s/p3t9nzbgfl216fursvzibpsgiby26hxh

Rawls, J. 1955, "Two Concepts of Rules", *The Philosophical Review*, 64, 3-32.

Schwyzer, H. 1969, "Rules and Practices", *The Philosophical Review*, 78, 451-67.

Searle, J.R. 1969, *Speech Acts: An Essay in the Philosophy of Language*, Oxford: Oxford University Press.

Williamson, T. 1996, "Knowing and Asserting", *The Philosophical Review*, 105, 489-523.

Wright von G.H. 1977, *Norm and Action: A Logical Enquiry*, London: Routledge & Kegan Paul; New York: Humanities Press.

The Ludic Background of Constitutive Rules in Bernard Suits

Filip Kobiela

University of Physical Education in Krakow

Abstract

The main purpose of the paper is to present and discuss Bernard Suits' account of constitutive rules presented in his *opus magnum—The Grasshopper. Games, Life and Utopia*—and in several minor contributions, which supplement or modify his original position. This account will be regarded as a crucial part of Suits' theory of ludic activities, mainly game-playing. The stress will be put on peculiarities of constitutive rules—their relation to ends in games, players' attitudes and their limitative nature. The analysis of the consequences of breaking a rule in different types of actions shows the essential difference between constitutive rules in games, and rules governing both technical activities, and non-game types of ludic activities. Because Suits' theory has been presented as an attack on Wittgenstein's claim concerning indefinability of games, this issue will be discussed as well.

Keywords: Bernard Suits, Ludwig Wittgenstein, Game-playing, Definition, Constitutive rules, Chess.

1. Introduction: Bernard Suits' Philosophy of Games

Bernard Suits is known mainly for his contribution to philosophical game theory. His legacy in this field consists of the seminal book *The Grasshopper. Game, Life and Utopia* (Suits 2014a) and several articles, being, among others, responses to criticisms, explanations and continuations of the 'grasshopperian' investigations.[1] The meaning of this legacy has been recognized in the circles of philosophers of sport. Outside this community, however, it is still awaiting wider recognition. It can be expected that this state of affairs will change due to the publication of the third, supplemented edition of *The Grasshopper* (2014a), and recent translations of this book (into Chinese, Japanese and Polish, 2016).

The main philosophical achievement of Suits is the formulation of an original and insightful definition of game-playing. The constitutive rules are one of

[1] List of Suits' articles on games and sport as well as articles on Suits might be found on the webpage of The International Association for the Philosophy of Sport: http://iaps.net/resources/suits/.

Argumenta 4,1 (2018): 125-137
ISSN 2465-2334

© 2018 University of Sassari

DOI 10.14275/2465-2334/20187.kob

the four elements of this definition, and play a crucial role in his theory of games. Ludic activities—mainly games, but also play and sport (the three considered in their mutual relationships constitute—to use Suits' idiom—"the tricky triad" (Suits 1988)) create the background of Suits' account of constitutive rules. The contrast: work-playing games, is the main conceptual vehicle for Suits' analysis, and his protagonist—Grasshopper from the ancient Aesop's fable—represents the ludic point of view and formulates a classical-type definition of game-playing. This characterizes the contemporary polemical context of the book—a discussion concerning the very possibility of the Grasshopper's endeavour—the definability of games.

2. The Discussion over the Definability of Games: Challenging Wittgenstein

In *Logical Investigations*, after introducing the term "language games" (*Sprachspiel*), Wittgenstein discusses the problem of the essence of a language-game, and hence of language in general. Instead of formulating a definition of language (which, according to Wittgenstein, should indicate something common to all that we call language) he emphasizes relationships which make us call all these phenomena "language". To illustrate this strategy Wittgenstein sketches an analysis designed to show that there is nothing in common to all proceedings that we call "games" (§66 and following). Instead of the classical notion of "essence", he offered the notion of "family resemblance" (*Familienähnlichkeit*) to indicate a network of overlapping and criss-crossing similarities among games:

> What is common to them all?—Don't say: "There *must* be something common, or they would not be called 'games'"—but *look and see* whether there is anything common to all.—For if you look at them you will not see something that is common to *all,* but similarities, relationships, and a whole series of them at that (Wittgenstein 1967: 31).

The "look and see" method consists in a review of series of activities called "games" (board-games, card-games, ball-games, noughts and crosses, patience, chess, tennis, ring-a-ring-a-roses) and noting retention and disappearance of certain characteristic features (amusement, wins and losses, competition, luck and skill).The purpose of the review is to show that there is nothing common to all the things called "games", but because of the network of resemblances among them, "'games' form a family" (Wittgenstein 1967: 32).

The analogy between *language games* and *games* discussed in §66 (which we may call here "ludic games") is designed to illustrate indefinability (in the classical sense) of language games, and the thesis concerning indefinability (in the classical sense) of ludic games, although clearly stated, is not a paramount consideration here. However, it is an interesting point in itself, and this very aspect of Wittgenstein's investigations draws attention of Bernard Suits.

Suits accused Wittgenstein of failing to follow his own advice "to look and see", and of deciding beforehand that games are indefinable (Suits 2014a: 1). Challenging Wittgenstein's claim Suits presented his own analysis that ends with the construction of a classical-type definition: there is a set of four necessary and sufficient conditions for being a game. If Suits is right in claiming that

games are definable, then efforts of defining other notions might be seen as more promising than shown in the light of Wittgenstein's anti-essentialism. The comparison between Wittgenstein's and Suits' accounts of games is thus also important beyond the philosophy of games.

3. The Means-end "Deduction" of Category of Game-playing

In the search for the definition of game-playing Suits engages the contrast between ant-type activities ("technical activity" or working) and grasshopper-type activities (playing games). Assuming that playing games is different from working, Suits is trying to find an aspect which differentiates working from playing games. He characterizes work as an "activity in which an agent seeks to employ the most efficient available means for reaching a desired goal". This in turn allows for formulating a very basic feature of games:

> Since games, too, evidently have goals, and since means are evidently employed for their attainment, the possibility suggests itself that games differ from technical activities in that the means employed in games are not the most efficient. Let us say, then, that games are goal-directed activities in which inefficient means are intentionally chosen (Suits 2014a: 24).

The crucial (and perhaps surprising) idea in the entertained possibility is that games are different from work not because they have different ends, but because the means that are used to achieve these ends are different, and more precisely—less efficient. This is sufficient to formulate short and simplified version of the definition: playing a game is "the voluntary attempt to overcome *unnecessary obstacles*" (Suits 2014a: 43).

Guided by the abovementioned idea, Suits offered a detailed analysis (that can be only roughly sketched here) of game-playing, which reveals four elements of the definition. The first element of this analysis is the *goal*, and the analysis shows the ambiguity of the term. A long distance runner might, for example, say that his goal is to:

1) win the foot race
2) cross the finish line ahead of the other contestants.

1 and 2 are not two different formulations of one goal, since it is possible to achieve 2 without achieving 1, but not vice versa. According to Suits, 1 presupposes 2, but 2 does not presuppose 1 and thus constitutes an elementary component of the foot race. In contrast, 1 is more complex, since it requires obeying the rules of a foot race (which, for example, forbids taking an underground during a marathon race, like in the famous case of Rosie Ruiz (Amdur 1980)). The kind of goal which is illustrated by 2 might be described generally as *a specific achievable state of affairs*.

> I suggest that this kind of goal be called the *prelusory* goal of a game, because it can be described before, or independently of, any game of which it may be, or come to be, a part. In contrast, winning can be described only in terms of the

> game in which it figures, and winning may accordingly be called the *lusory* goal of a game (Suits 2014a: 38-39).[2]

In a subsequent explanation of the game-independence of the prelusory goal, Suits claims that the term "identified" is more appropriate than the term "described", because in some cases (e.g. chess) the goal might be more conveniently displayed rather than described (Suits 2014a: 38-39).

It might seem that the prelusory goal is a brute fact (in Searle's sense), whereas the lusory goal is a social or institutional fact, however—what is clearly visible in case of chess—the prelusory goal might also be an institutional fact. The crucial difference between 1 and 2 lies in the fact that the prelusory goal is "distilled"[3] not from any institutional aspects, but from any reference to the way of achieving it: "By omitting to say *how* the state of affairs in question is to be brought about, it avoids confusion between this goal and the goal of winning" (Suits 2014a: 38). Suits illustrates this distinction by many examples, apart from the abovementioned foot races, he discusses, for example, mountain climbing: its prelusory goal is to be on the top of the mountain, which can be achieved by means of using a helicopter, whereas its lusory goal is *to climb* the mountain (Suits 2014a: 90-92).

Since games are goal-directed activities, some means must be used in efforts undertaken to achieve a (pre)lusory goal, hence there is the second element of game-playing—the means. Like goals, means are divided into prelusory and lusory ones, and because the lusory means are means which are permitted by the constitutive rules, the third element of Suits' analysis emerges—the constitutive rules.

The rules in games (like goals and means) are also of two kinds. Rules responsible for defining permitted (legal) means in achieving the prelusory goal Suits calls "constitutive rules". They are associated with the prelusory goal, and with it become sufficient for characterizing the nature of a given game. The mechanism of dividing the possible means into legal and illegal is based on the fact that constitutive rules are proscriptions of certain means useful in achieving prelusory goals:

> think of any game at random. Now identify its prelusory goal: breasting a tape, felling an opponent, or whatever. I think you will agree that the simplest, easiest, and most direct approach to achieving such a goal is always ruled out in favour of a more complex, more difficult, and more indirect approach (Suits 2014a: 40).

Thus the nature of constitutive rules in games lies in the restrictions imposed on the means: "We may therefore define constitutive rules as rules which prohibit use of the most efficient means for reaching a prelusory goal" (Suits 2014a: 40). The other kind of rules in games—associated with lusory goals—is called the

[2] Suits entertains also but finally rejects a different terminology: "It is tempting to call what I have called the pre-lusory goal the goal in a game and the lusory goal the goal of a game, but the practice of philosophers like J.L. Austin has, I believe, sufficiently illustrated the hazards of trying to make propositions to carry a load of meaning which can much better be borne by adjectives and nouns" (Suits 2014b: 21). Part of this article was later used in *The Grasshopper*, however this remark has been removed by Suits.

[3] Suits (2006: 5) in the context of prelusory goal uses the term "isolating" (elements of playing games for examination).

rules of skill. In contrast to the proscriptive nature of constitutive rules, the rules of skill are instructions suggesting the most efficient way of achieving a prelusory goal. There is an interesting balance between these two types of rules: whilst constitutive rules proscribe the use of the most efficient means, rules of skills prescribe the most efficient means, however only within the frame defined by the constitutive rules. Games appear to be paradoxical activities, since they are governed by both, the principle of inefficiency: (its constitutive rules prohibit using the most efficient means of achieving the prelusory goal, making this type of activities different from technical activities), and by the principle of efficiency (rules of skill prompt using the most efficient—inside the legal area—means of achieving the lusory goal).

Finally, the last element of the analysis (and the definition of game-playing) deals with the reason for acceptance of the constitutive rules. In contrast to the extra-lusory activities, in which acceptance of rules limiting the most efficient ways of achieving the desired goal is imposed by some external force, in games—as they are voluntary activities—this acceptance flows from players' free decisions based on understanding that the acceptance of the rules is a necessary condition for the game to exist.[4] Suits calls this attitude *lusory attitude* and defines it as the "acceptance of constitutive rules just so the activity made possible by such acceptance can occur" (Suits 2014a: 43).[5] In other words, in games "the sole reason for accepting the limiting rule is to make possible such activity" (Suits 2014a: 72).

Suits' analysis culminates in the famous formula:

> To play a game is to attempt to achieve a specific state of affairs [prelusory goal], using only means permitted by rules [lusory means], where the rules prohibit use of more efficient in favour of less efficient means [constitutive rules], and where the rules are accepted just because they make possible such activity [lusory attitude] (Suits 2014a: 43).[6]

The formula presents a set of necessary and sufficient conditions of being a certain activity: *game playing,* and as such does not provide a definition of a *game* (as the product of this activity). The following definition paraphrases Suits' original definition into explicit definition whose definiendum contains only the term *games*:

[4] The acceptance of the rules as a necessary condition of there being a game has been recognized by Caillois (2001: 7) "The game is ruined by the nihilist who denounces the rules as absurd and conventional, who refuses to play because the game is meaningless".

[5] This formulation is not far away from Caillois' (2001: 29) psychological observation: "the pleasure experienced in solving a problem arbitrarily designed for this purpose also intervenes, so that reaching a solution has no other goal than personal satisfaction for its own sake".

[6] The definition of game playing was first presented in Suits 1967. Suits offers also an auto-ironic caricature of his own definition:"a game is when, although you can avoid doing something disagreeable without suffering any loss or inconvenience, you go ahead and do it anyway" (Suits 2014a: 56). This is one of many examples of sense of humour presented in *The Grasshopper*. The definition of game playing was first presented in Suits 1967.

> To put matters a little more precisely […] games are rule-governed activities in which (a) a participant pursues a prelusory goal, (b) using only those means permitted by the rules, (c) where those rules exclude more efficient in favor of less efficient means of realizing the prelusory goal, and (d) in which the participant accepts the limitations to make the activity possible (Berman 2013: 1).[7]

Constitutive rules appear in Suits' analysis together with the hypothesis of the means' inefficiency in games. From that point onwards, constitutive rules become one of the most important topics of Suits' investigations, and his account of constitutive rules in games seen as the limitations of permitted means to achieve an end is the most characteristic aspect of his theory.

4. Definition of Game-playing: Objections and Defences

The definition [mentioned above] seems plausible but also thought provoking, or rather objection-provoking, and has not gone unchallenged by Suits' critics.[8] However, the most interesting discussions have been presented by Suits himself. After presenting the above sketched construction of a definition (in Chapter III of *The Grasshopper*), Suits presented a series of attacks on the definition (launched by the aptly named ant Scepticus) and defences formulated by Suits' protagonist—Grasshopper. The main task of this discussion is to deepen the understanding of the definition by entertaining several objections of two kinds: 1) the definition is too narrow (error of exclusion) and 2) the definition is too broad (error of inclusion).[9]

Appendix one to Suits' book contains a different kind of objection, which deserves special attention: some activities (e.g. foot races) that satisfy Suits' definition of games are not even *called* games. From the viewpoint of analytical philosophy of language, Suits' response to this objection is one of the most important parts of this polemics. His analysis might be classified as a contribution to the philosophy of constructing definitions. Suits starts with the observation that Wittgenstein's question, whether all things *called* games have something in common, is very different from the question whether all things that *are* games have something in common. This leads to a distinction between things that are *called* games, and things that *are* games, and in consequence, to discussion of relationship between properties of *being called* a "game" from *being* a game.

Suits claims that Wittgenstein erroneously presumes that according to essentialists all things called by the same name have the same definition. According to Suits, the pattern of linguistic phenomena (naming) does not necessarily fit the pattern of definitions (being). This is because resemblance between things is based on the directly observable behavioural properties. To find a real differ-

[7] About some complication of such manipulation with the definition see Berman 2015: 1, 2.
[8] Perhaps the most surprising feature of the definition is that it does not mention competition, and this issue has been raised by King 2015. However, this objection might be rebutted by indicating how elements of the definition involve competition (cf. Vossen 2004). For defence of Wittgenstein see Ellis 2011.
[9] One objection (put forward by McBride 1979) uses Wittgenstein's own invention: game (or rather alleged game) "Sun Earth and Moon" (SEM) described in N. Malcolm's memoirs. Suits' response to this objection has been added to the second edition of *The Grasshopper* as "Appendix Two: Wittgenstein in the meadow" (Suits 2014a: 211-16).

ence between classes of things (to formulate a proper definition) one should take account not just of superficial properties, but also of some more fundamental features.[10] Suits' definition is an example of capturing such a hidden mechanism (in which constitutive rules play a crucial role) that lies behind the directly observable behavioural properties. On that ground Suits claims that a proper criterion for being a game is not being called a game (which might be misleading, as in the case of ring-ring-a-roses) but meeting the definition of game playing. Some activities, such as foot races, meet the definition, but are not called games. Our linguistic habits might be sometimes irregular, and thus the theory based on definitions should have the priority in taxonomy over the ordinary language.

5. Pre-lusory Goal in Chess, Double Function of Constitutive Rules and Lusory Institution

Suits belongs to the group of philosophers who pay special attention to chess as representing the functioning of constitutive rules. *The Grasshopper* introduces a chess-related analysis as a response to an objection against the notion of the prelusory goal. The objection states that the offered definition is too narrow, since it cannot capture games like chess. The objection might be structured as follows:

1: "the *prelusory* goal of a game […] can be described before, or independently of, any game of which it may be, or come to be, a part" (Suits' definition of prelusory goal);
2: "the end in chess is […] to place your pieces on the board in such an arrangement that the opponent's king is, in terms of the rules of chess, immobilized"[11] (Suits' definition of "prelusory" checkmate);
3: "alleged prelusory goal of chess is already saturated with rules and is therefore not a prelusory goal as defined" (consequence of 2);
Conclusion: since chess is undoubtedly a game and it does not satisfy the definition in question (because of the absence of the prelusory goal in chess), *the definition is too narrow*.

Suits' answer to this apparently convincing objection consists in indicating two different ways of functioning of constitutive rules in games. In the first way—called descriptive—the constitutive rules "are used to *describe* a state of affairs"—for example "it is necessary to refer to the rules of chess in describing checkmate" (Suits 2014a: 48). The descriptive use of constitutive rules creates "an *institution* of chess which can be distinguished from any individual game of chess". In the second way—called prescriptive—the constitutive rules are used to *prescribe* a procedure. It is thus possible to achieve a descriptive checkmate in chess without playing a *game* of chess. Now, the objection follows form the confusion of these two ways. Suits is arguing that in chess there is an end analytically dis-

[10] To use the paradigmatic example: consider the expression *sunrise*. According to the modern astronomy, this phenomenon should be rather called 'earth-dip'. As this case clearly shows, wrong names often occur due to cultural lag (Suits 2014a: 202-203).
[11] Suits formulation is equivalent to the official FIDE definition: "The objective of each player is to place the opponent's king 'under attack' in such a way that the opponent has no legal move", *Laws of chess*, article 1.2.

tinct from winning, and that this end might be achieved without following the rules of chess.[12]

> Now, without going outside chess we may say that the means for bringing about this state of affairs consist in moving the chess pieces. The rules of chess, of course, state how the pieces may be moved; they distinguish between legal and illegal moves. Since the knight, for example, is permitted to move in only a highly restricted manner, it is clear that the permitted means for moving the knight are of less scope than the possible means for moving him. It should not be objected at this point that other means for moving the knight—e.g., along the diagonals—are not really possible on the grounds that such use of the knight would break a rule and thus not be a means to winning. For the present point is not that such use of the knight would be a means to winning, but that it would be a possible (though not permissible) way in which to move the knight so that he would, for example, come to occupy a square so that, according to the rules of chess, the king would be immobilized (Suits 2014a: 35).

The example shows the possibility of achieving the prelusory goal of chess without achieving its lusory goal. Of course a player who made such a move would not be, in a strict sense, playing chess, but would rather be cheating at chess. But the same situation occurs in all games and chess are games "because of an 'arbitrary' restriction of means permitted in pursuit of an end" (Suits 2014a: 36).

In the context of different ways of using constitutive rules pre-lusory goal of a game might be thus called "descriptive goal", whereas lusory goal of a game might be called "prescriptive goal". After these considerations we are in a position to supplement the original definition of prelusory goal in a way which will prevent it from this type of objection: a prelusory goal of a game can be described before, or independently of, any individual game *(understood as a prescriptive use of rules)* of which it may be, or come to be, a part, *but not necessarily independently of an institution of a game, defined by the descriptive use of the rules of the game.*

The status of chess and its goal is a recurring topic in Suits-related literature,[13] and in a defence of his position Suits took the opportunity to expand his definition of the institution of game (or the lusory institution):

> The institution of game *x* is a body of diverse meanings and practices that have in common the fact that all derive their being (that is, they are what they are) by being related, in one way or another, to game *x* (Suits 2006: 4).

The distinction between individual games and the institution of game is, according to Suits, presupposed by the identification of certain types of behaviour associated with playing games. A model attitude is a (genuine) player, who is trying to achieve a prelusory goal of the game obeying its constitutive rules. The

[12] This distinction has been recognized by Żełaniec (2013: 148), who observed that official (FIDE) chess definition of winning is "neither extensionally nor intensionally the same as 'winning'. Had FIDE not expressly defined chess as a game, the 'placing the opponent's king' etc. could have been practiced in a non-game-like fashion, for instance, as a religious rite".

[13] Butcher, Schneider 1997 contains an attack on Suits' notion of prelusory goal in which chess is discussed; Vossen 2008, also, raises this issue.

remaining attitudes characterize rather *quasi*-players than genuine players. A cheat (who plays a special role in the problem of the so-called formalism—see below) wants to achieve a condition which is, descriptively, a condition of a goal, but he violates the rules of the game in his efforts to do so. A cheat operates within the institution of the game, because

> he violates the rules in their prescriptive application only because of his expectation that they will be observed in their descriptive application […]. In terms of their dependence upon institutions, cheaters at games are precisely like liars in everyday life (Suits 2014a: 50).[14]

While cheating is a relatively frequent attitude towards the rules and goals in games, the opposite attitude—that of a trifler—is less known. Unlike the cheat (and a genuine player), the trifler is a player whose actions, although all legal, are not directed to achieving the goal of the game. He is not "playing to win", although he is not necessarily "playing to lose". The trifler only recognizes the rules (he is only using lusory means), but not the goal, and whilst he is operating within the institution of game, he is in fact not playing a game (Suits 2014a: 49-51).[15]

The case of the trifler is interesting because his behaviour proves that merely acting in accordance with the rules is not sufficient for playing a game (cf. Lorini 2012: 141-42). The trifler "lacks zeal in seeking to achieve the prelusory goal" (Suits 2014a: 50) of the game, but because he is not breaking any rule of the game, his behaviour might be properly characterized only on the meta-level. "Play to win" does not belong to the set of rules of the game, but perhaps might be characterized as a meta-rule.

6. Formalism and Logical Incompatibility Thesis

It is relatively common practise to break a rule during playing games. However, a question arises if a person breaking the rules (a "cheater") could really win the game, since winning is a part of playing, and playing presupposes obeying the rules. One might then ask, for example, if the goal scored illegally by Maradona and known as "the *hand* of God" was *really* a part of the game, whose rules forbid attackers to use *hands*? Suits' theory of game playing belongs to the family of theories called "formalism";[16] its roots might be find as early as in Huizinga: "The rules of a game are absolutely binding and allow no doubt […] as soon as the rules are transgressed the whole play-world collapses" (Huizinga 1949: 11). According to Suits, "Rules in games thus seem to be in some sense inseparable from ends, for to break a game rule is to render impossible the attainment of an end" (Suits 2014a: 26). In technical actions is possible to break a rule without destroying the original end of the action (e.g. get to an appointment without obeying the traffic lights), but in games "If the rules are broken the original end becomes impossible of attainment, since one cannot (really) win the game unless

[14] A similar example is used by Żełaniec 2013: 101 "when we move a bishop non-diagonally we offend against a rule of chess in quite a similar way as when we offend against a rule of human communication when we lie".

[15] The last type of quasi-players is spoilsport, who acknowledges neither rules nor goal.

[16] However, some philosophers argue that the concept of lusory attitude admits a different interpretation of Suits (cf. Frias 2016: 58).

one plays it, and one cannot (really) play the game unless one obeys the rules of the game (Suits 2014a: 26). The argument—central for formalism—is known as "logical incompatibility thesis" (Morgan 1987) since it is logically impossible at the same time to obey the rules (by the definition of game-playing) and not to obey the rules (by the definition of cheating).

Formalism emphasizes the problem of breaking a rule. The following remark by Suits compares consequences of breaking different kinds of rules in games: "To break a rule of skill is usually to fail, at least to that extent, to play the game well, but to break a constitutive rule is to fail (at least in that respect) to play the game at all" (Suits 2014a: 40). Because the rules of skills are instances of regulative rules, the remark allows the following formulation of a criterion to distinguish between a constitutive and non-constitutive (regulative) rule:

> *A given rule is a constitutive one (for a given activity) if and only if breaking this rule makes logically impossible attaining the end of this activity.*

On this ground the new rules introduced to games (or modifications of old rules) should be understood not as regulative, but rather constitutive rules. One might say that a rule modifying an already existing game (like, for example, the prohibition of catching the ball by the goalkeeper if the ball is passed by the player of his own team) satisfies Searle's definition of a regulative rule (it regulates an antecedently existing form of behaviour, activity whose existence is independent of the rule) (Searle 1964: 55). However, on the ground of the rule-breaking criterion, the rule is a constitutive one, and its violation is not just a mistake in the art, but it results (in a properly refereed game) in an immediate stopping of the game and imposing punishment. This criterion leads us to the question of game-status of some sport disciplines.

7. Games, Performative Sports and Constitutive Rules

The restrictive aspect of Suits' understanding of constitutive rules is clearly visible in the context of comparison between the two different types of sport activities distinguished by Suits. Originally Suits classified all sports as games (in his meaning of the term); *The Grasshopper* is full of examples of games that are at the same time sports: foot-races, boxing, high-jump, golf, badminton, hockey, baseball, football, car racing etc. However, later, Suits revised this view by claiming that some sports (like diving and gymnastics competitions) are not games in his meaning of the term, and calls this secondary group of sports "performances", "performative sports" or "judged events",[17] whereas the main group of sports being games (in his meaning of the term) called "refereed events" or simply "games". Suits does not deny the fact that sports which are not games are rule-governed, although the nature and way of functioning of the rules in the games and performative sports are very different. Reformulating the short version of his definition of games from *The Grasshopper*, Suits claims that the essence of games (and, obviously, all sports that are games) lies in "erecting artificial constraints just so those constraints can be overcome". In this type of activities constitutive rules are crucial, because they define which actions are legal, and breaking such a rule results in penalties imposed by referees. The nature of performative sports is quite different:

[17] In the contemporary literature this group is called "judged sports" (see Hurka 2015).

> Now it may be objected that, contrary to what I have said, there clearly *are* rules that must be followed while actually engaged in performative sports. For example, the gymnast must not falter or stumble after dismounting from the parallel bars. It is perfectly permissible to call such a requirement a rule, but it is quite clear, I should think, that such rules are entirely different from, say, the offside rules in football and hockey offside rule is what has come to be called, by me and many others, a constitutive rule, while the standard of a clean dismount from the parallel bars is a rule of skill, or a tactical rule, or a rule of practice. […] Now, with the exception of what I have called pre-event rules (the steroid example), the rules to which the judges of performances address themselves are, I submit, rules of skill rather than constitutive rules (Suits 1988: 5-6).

The essence of performative sports lies in postulating ideals to be approximated, and the jury's task is to evaluate the proximity of a given performance to this ideal. In the course of the performative sport there are no constitutive rules (!) the breaking of which might be punished. The rules that define ideals evaluated by the jury (and which, also, create skills specific for a given performative sport) are, according to Suits, not constitutive rules, since they do not prescribe any behaviour. Obviously, every case of breaking such rules results in deducing the points by the jury, but it should be understood not as a failure at playing a sport at all (due to the logical incompatibility thesis), but only as a failure at playing good.

The distinction between games and performative sports clearly shows that Suits' account of constitutive rules is narrow: because the essence of these rules consists in imposing limitation, restriction or prohibition, they should rather be called constitutive-restrictive rules, constitutive-limitative rules, or constitutive-prohibitory rules.

8. Conclusion

Suits claims that Wittgenstein exhibits "idolatrous insensitivity to definitions" and classifies the idea responsible for this induced insensitivity using Bacon's idiom: the idol of family resemblance. Since a standard user of an ordinary language is a working essentialist, this idol cannot be classified as an idol of the Market Place, and thus it is called an Idol of Academy (Suits 2014a: 209). In contraposition to Wittgenstein Suits claims that games are definable, and that constitutive rules (in constitutive-restrictive sense) are a part of their definition. Even defenders of the Wittgenstein's "family" account of games must admit that his "look and see" method ignores highly probable candidates for essential features of all games, like being *goal directed*, *voluntary* (point emphasized by both Huizinga and Caillois), and—most importantly—*rule-constituted*, institution-based activities. The latter omission is especially surprising since Wittgenstein himself devoted a lot of considerations to the problem of relation between rules and games,[18] and his analogy between language games and "ludic" games seems to be based on the crucial function of rules in both elements of the analogy. Although Suits' account of games and constitutive rules might be not the "last word" in the field, it deserves a serious debate since it offers some important insights into the constitution of the ludic aspects of a social being. I believe that

[18] Wittgenstein even used the term "constitute" in this context (cf. Conte 1998: 252).

Suits' definition of game playing does not only effectively challenge Wittgenstein's position on games, but also encourages a search for definitions in general, hence its significance goes beyond the boundaries of the philosophy of games.

References

Amdur, N. 1980, "Who is Rosie Ruiz", *New York Times*, April 2.

Berman, M.N. 2013, "Sprints, Sports, and Suits", *Journal of the Philosophy of Sport*, 40, 1, 163-76.

Berman, M.N. 2015, *Sport as a Thick Cluster Concept* (conference draft), https://www.humanities.utoronto.ca/uploaded_files/content/1630/file/Berman_Cluster_concept_of_sport.pdf (retrieved 20.II.2017).

Butcher, R.B and Schneider, A.B. 1997, "Pre-lusory Goals for Games: A Gambit Declined", *Journal of the Philosophy of Sport*, 24, 38-46.

Caillois, R. 2001, *Man, Play and Games*, New York: Free Press of Glencoe.

Conte, A.G. 1988, "Eidos. An Essay on Constitutive Rules", in Di Bernardo, G. (ed.), *Normative structures of the social world*, Dordrecht: Kluwer, 251-57.

Ellis, J. 2011, "On the Concept of a Game", *Philosophical Investigations,* 34, 4, 381-92.

Frias, J.L. 2016, "The Reconstructive and Normative Aspects of Bernard Suits's Utopia", *Reason Papers*, 38, 1, 51-64.

Huizinga, J. 1949, *Homo Ludens. A Study of the Play-element in Culture*, London: Routledge & Kegan Paul.

Hurka, T. 2015, "On Judged Sports", *Journal of the Philosophy of Sport*, 42, 3, 317-25.

King, P. 2015, "Games and Pastimes", Conference draft, https://www.humanities.utoronto.ca/uploaded_files/content/1630/file/King_games_and_pastimes.pdf (retrieved 20.II.2017).

Lorini, G. 2012, "Can Constitutive Rules Create a Practice?", *Praxis Filosofica*, 34, 139-48.

McBride, F. 1979, "A Critique of Mr. Suits' Definition of Game Playing", *Journal of the Philosophy of Sport*, 8, 59-65.

Morgan, W. 1987, "The Logical Incompatibility Thesis and Rules", *Journal of the Philosophy of Sport*, 14, 1-20.

Searle, J. 1964, "How to Derive 'Ought' from 'Is'", *The Philosophical Review*, 73, 1, 43-58.

Suits, B. 1967, "What is a Game", *Philosophy of Science*, 34, 2, 148-56.

Suits, B. 1988, "Tricky Triad: Games, Play and Sport", *Journal of the Philosophy of Sport*, 15, 1-9.

Suits, B. 2006, "Games and their Institutions in Grasshopper", *Journal of the Philosophy of Sport*, 33, 1-8.

Suits, B. 2014a, *The Grasshopper: Games, Life and Utopia*, third edition, Peterborough: Broadview Press.

Suits, B. 2014b, "The Elements of Sport", in Holt, J. (ed.), *The Philosophy of Sport: Core Readings,* Peterborough: Broadview Press, 19-33.

Vossen, D. 2004, "The Nature and Classification of Games", *Avante*, 10, 1, 53-68.

Vossen, D. 2008, "Chess Is Not a Game", in Hale, B. (ed.), *Philosophy Looks at Chess*, Chicago: Open Court, 191-208.

Laws of chess, World Chess Federation/Federation internationale des echecs (FIDE), http://www.fide.com/component/handbook/?id=124& (retrieved 20.II.2017).

Wittgenstein, L. 1967, *Philosophical Investigations*, 3rd ed., Oxford: Blackwell.

Żełaniec, W. 2013, *Create to Rule. Studies on Constitutive Rules*, Milano: LED Edizioni Universitarie.

Constitutive Rules and the Internal Point of View

Corrado Roversi

University of Bologna

Abstract

In this paper, I connect J.R. Searle's concept of constitutive rules and H.L.A. Hart's concept of internal point of view and look for an extension of this joint paradigm in institutional ontology. I make a distinction between five different perspectives about an institution—structural, teleological, axiological, strategic, and sociological—and connect these perspectives to three kinds of concepts: institutional, meta-institutional, and para-institutional. In the light of these distinctions, I submit that an explanation of institutional phenomena requires a three-dimensional ontology consisting of a structure (framed by constitutive rules), a conceptual background, and an actual practice. I then proceed by showing that this three-dimensional ontology makes it possible to specify Hart's famous distinction between internal and external point of view (the latter being either moderate or extreme) into a more shaded distinction between six different approaches to an institution, exemplified by six different archetypical characters.

Keywords: Constitutive Rules, Institutions, Social Ontology, Internal Point of View, H.L.A. Hart.

1. Introduction: H.L.A. Hart and J.R. Searle

When considering the main authors who have been giving a major contribution to the philosophical discussion on constitutive rules, the role of H.L.A. Hart is often overlooked. This is quite odd, in a sense, given the focus he devotes to the enabling character of several legal norms, as opposed to the disabling character of duty-imposing commands. He writes in this regard:

> The criminal law is something which we either obey or disobey and what its rules require is spoken of as a 'duty.' [...] But there are important classes of law where this analogy with orders backed by threats altogether fails, since they perform a quite different social function. Legal rules defining the ways in which valid contracts or wills or marriages are made do not [...] impose duties or obligations. Instead, they provide individuals with facilities for realizing their wishes. [S]ome of the distinctive features of a legal system lie in the provision it makes, by rules of this type, for the exercise of private and public legal powers. If such

Argumenta 4,1 (2018): 139-156
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.rov

> rules of this distinctive kind did not exist we should lack some of the most familiar concepts of social life, since these logically presuppose the existence of such rules (Hart 1994: 27-28, 32).

Hart speaks here of concepts that logically presuppose rules and are connected to private and public powers. He calls these rules secondary, as opposed to primary rules of obligation.

Another important distinction we find in Hart's philosophy of law is that between an internal and an external point of view on rules. Hart maintains that one of the crucial features of the social phenomenon we call "existence of a rule" is that people have an *internal* point of view, namely, they treat the rules as standards for justifying their behaviour and possibly criticizing that of others. On the other hand, when people simply describe regularities of behaviour or the beliefs of others about their own rules, they adopt an *external* point of view. Usually, people who adopt an internal point of view use a specific normative vocabulary, which in the case of the primary rules of obligation involves expressions like "You ought," "You must," "That is right," and "That is wrong" (see Hart 1994, 55ff.).

Now, in Hart's view, the existence of a legal system requires an internal point of view not only on primary rules of obligation but also on secondary rules. When officials have an internal point of view in regard to secondary rules, their normative vocabulary is enriched by a set of concepts that is much broader than the one which includes only simple deontic qualifications of behaviour:

> With the addition to the system of secondary rules, the range of what is said and done from the internal point of view is much extended and diversified. With this extension comes a *whole set of new concepts and they demand a reference to the internal point of view for their analysis*. These include the notions of legislation, jurisdiction, validity, and, generally, of legal powers, private and public (Hart 1994: 98-99; italics added on last occurrence).

This phenomenon of rule-created concepts connected with an internal point of view, namely, an attitude of general acceptance of the rules, is similar to the one that John R. Searle describes in terms of constitutive rules. In his theory, Searle shows how human beings create institutional phenomena by collectively accepting constitutive rules assigning status functions to "brute" facts. In Searle's theory, as in Hart's, the acceptance of constitutive rules relates to a specific internal perspective in which agents act according to the rules and use the concepts defined by them: This is the perspective where agents interact with the institutional facts made possible by those rules. The institutional world made possible by constitutive rules so understood is pervasive, and to appreciate as much we can look at the following passage by Searle:

> One of the advantages of living in other cultures is that one can become more acutely conscious of the different and unfamiliar institutional structures. But at home one is less aware of the sea of institutionality. I get up in the morning in a house jointly *owned* by me and my *wife*. I drive to do my *job* on the *campus* in a car that is *registered* to both of us, and I can drive *legally* only because I am the holder of a *valid California driver's license*. On the way, I *illegally* answer a cell phone call from an old *friend*. Once I am in my *office* the weight of institutional reality increases. I am in the *Philosophy Department* of the *University of California* in

> *Berkeley*. I am surrounded by *students*, *colleagues*, and *university employees*. I teach *university courses* and make various *assignments* to my *students*. The *university pays* me, but I never see any *cash* because my *pay* is *deposited* automatically into my *bank account*. […] All the italicized expressions in the previous paragraph refer to institutional reality in its various aspects. Institutional facts range all the way from the informality of friendship to the extreme legal complexities of international corporations (Searle 2010: 90-91).

For all the similarities between Searle's concept of constitutive rules and Hart's concept of secondary rules, the two are not identical. In particular, Hart's secondary rules are meta-rules, that is, rules that define concepts and institutions related to the way in which other rules in the legal system must be dealt with (see Hart 1994, 96-98). By contrast, Searle's constitutive rules simply attribute status functions that are not necessarily connected with the use of other institutional rules. Hence, one could conclude that Hart's secondary rules are a subset of Searle's constitutive rules. However, it cannot be denied that both Hart's and Searle's theories seek to address the way rules can provide us with new ways of thinking and acting.

The question, then, is: can Hart's internal point of view on an institution and Searle's concept of constitutive rules be sufficient to understand the conceptual domain linked to that institution and the behaviour of the agents who interact with it? I submit that they cannot: other elements are needed. If we are to understand the "life" of an institution, we need to be acquainted not only with its specific rule-constituted concepts but also with the social purpose and meaning of the overall practice these constitutive rules are part of: we have to view constitutive rules in light of a background practice that the institution instantiates. And this in turn gives rise to a more nuanced distinction between the possible perspectives that agents may have on an institution. It is not only a matter of whether we take an internal or an external point of view on rule-defined facts, or whether we accept them or not, or how we interact with them: it is also a matter of how the institution's structure, purpose, and core values are conceived, how we act strategically with respect to it, how interested we are in its typical sociological outcomes, and so on. Hence, the kind of conceptual broadening that Hart and Searle rightly call for in explaining law and institutional facts in general requires a knowledge not only of constitutive rules but also of the background in which these rules are embedded, a background that moulds the possible kinds of perspectives we can take when dealing with institutions and living in an institutional setting. In this sense, we have to push Hart and Searle's perspective further.

In this paper, I present five different kinds of perspectives on rule-constituted institutions—structural, teleological, axiological, strategic, and sociological—illustrating how these perspectives are typically connected with different kinds of concepts. I also argue that not all these concepts can be explained in terms of constitutive rules: these rules only form the *structure* of institutions, but a full explanation of institutional concepts needs to be complemented by bringing in both the background and the practice connected to that structure. Further, I show how the distinction between different kinds of perspectives on an institution makes it possible to nuance and further specify the distinction between the internal and external points of view put forward by Hart. The structure of the paper is as follows. In Section 2, I present the *structural* perspective, arguing that

it typically focuses on *institutional* concepts, namely, concepts constituted through rules. In Sections 3 and 4, I respectively deal with the *teleological* and *axiological* perspectives, and show how these are based on *meta*-institutional concepts of two different kinds. In Section 5 and 6, I introduce the *strategic* and *sociological* perspectives and connect them to *para*-institutional concepts, again of two different kinds. In Section 7, I relate these perspectives to Hart's distinction between the internal and external points of view, showing how they can specify and enrich that distinction. Finally, in Section 8, I draw some conclusions.

2. The Structural Perspective

Suppose I want to teach my daughter Adriana to play chess. I lay the chessboard on the table and look at her: she is curious, she really wants me to teach her what this strange thing is. But I immediately face a crucial question: How should I teach her? I begin to ponder several options but am pressed by her insistence. So I decide to adopt this method: "Ok, Adriana, this is the chessboard. And this is the king, that is the queen, and here we have the two bishops, the rooks, the knights, and all the pawns. Now, let me explain how chess works: the king can move only by one square, whereas the queen can move in any direction .... Oh, and any piece can capture any other piece simply by landing on the same square once its own move has been completed ...."

Here I would be taking a structural perspective on the game of chess. Indeed, what I am describing to Adriana is the *structure* of the game, and in particular that of its basic concepts: the pieces, the mechanics by which pieces can be "taken" or captured, the possible moves. I am teaching her the game's constitutive rules. According to Searle, all these rules can be traced to a common cognitive and linguistic process, that is, they are "standing declarations" that attribute status functions connected to deontic powers (see Searle 2010: 101ff.), and they make it possible to qualify some facts and actions as institutional. In this way, my act of taking a piece of material shaped in such and such a way and moving it across a checkered board becomes my moving a *bishop* on a *chessboard*.

In Searle's view, not all institutional facts depend on constitutive rules, but in general institutions do. In *Making the Social World*, he distinguishes between three different types of institutional facts (Searle 2010: 94-98).[1] The first type he calls "institutional fact[s] without an institution" (Searle 2010: 94). Here, by collective acceptance the members of a community simply assign status functions and deontic powers to a given entity, as when someone is collectively recognized as king. In this case, while a status function and some deontic powers are attributed to a concrete and specific entity, there is no standard connection between entities *of a given kind* and the relative status function: what matters is the *individual* entity.

In Searle's view, institutions, and hence institutional facts of the second kind come into being when we do have this standard and regular connection between kinds of entities and status functions. Here we have status-function attributions on the basis of constitutive rules, as when the members of a community "evolve a standard procedure for selecting the king" (Searle 2010: 96). The cru-

[1] I will skip Searle's third type of institutional fact, because as far as we are concerned it does not differ significantly from the second.

cial difference with respect to institutional facts of the first kind is that here the rule has to do not with a specific individual recognized as the king, but rather with a procedure on which basis *any king* can be selected *on any occasion*.

Now, typically, when constitutive rules come in systems (as in the game of chess) they create new concepts, because the attribution of a status function connected to deontic powers is relative to the other elements of the systems, and can have no meaning unless it is connected with them. Hence, the meaning of an institutional term depends on the relations between the elements of the system and is typically new, or at least it is a substantial modification of another, pre-existing concept (this is the sense in which, for example, the concept of a bishop in chess is created by the rules of chess). This is the main feature of what I will be calling *institutional concepts*, which is that constitutive rules are necessary and sufficient conditions for the existence of such concepts because they create them.

In a sense, the structural perspective is a kind of internal point of view in Hart's sense. But it is a very narrow point of view, because in taking that perspective we limit our discussion to institutional concepts. When we reason within a purely structural perspective, the "internality" of our point of view is so extreme that we simply speak in technical terms, focusing on the system of rules and on rule-constituted concepts: we do not address questions like what is the overall point of the institution in question, whether it is just, or whether and how it is practiced. Law has an abundance of institutional concepts, and jurists and legal professionals can very well stick to a purely structural perspective. In a strictly legal-positivistic sense, they adopt the point of view of simple "norm technicians", analysing and combining normative propositional contents to establish the relative institutional structures and accurately determine the normative consequences that follow from them.

3. The Teleological Perspective

But let us go back to Adriana: she is not going to understand my explanation of chess. At first she will gaze at me in amusement, listening to my elaborate description of the rules and pieces, but soon she will grow impatient. The problem is not that she cannot understand what I am saying: she is smart. She now starts to move the pieces on the chessboard according to the rules, and does so correctly. We take a few turns making moves but then she gets bored. "What's the matter, Adriana?" I ask. "Aren't you enjoying yourself?" "Not at all," she retorts: "This is boring." "Why do you think this is boring? You are very good at moving the pieces." "Yes, but I don't understand *why* I have to move them!"

As mentioned, Adriana is smart. She realizes that the structural perspective is severely limited and cannot by itself provide a full picture of the game of chess. If Hart's concept of the internal point of view were confined within the structural perspective, it would be extremely wanting as an explanation of law: any Martian observer trying to understand the practice of chess would be in the same position as Adriana. What, then, should I say to her? The answer is quite simple: "Adriana, this is a game, and you should try to win!"

Here we have a perspective on the game of chess that is broader than the skeletal structural perspective. If we conceive the elements of chess as pieces with which we have to interact in order to win, we broaden our view beyond the strictly internal perspective where only the structural connections among ele-

ments are considered: we take the more embracive internal perspective, where those connections are understood to form part of a means-end structure—a teleological structure. What we are essentially doing in taking this teleological perspective is that we are viewing the constitutive rules of a given institution, along with the corresponding system of institutional concepts, against the background of a broader practice: in the case of chess, the broader practice of competitive game-playing, of which the game is an instance.[2]

As in the case of the structural perspective, the teleological perspective is connected with concepts of a kind that is relevant to the institution at hand. But these concepts cannot be described as institutional. In fact, going back to the example of chess, victory is not a rule-constituted institutional concept on the same level as, say, "castling" or "checkmating".[3] Whereas the conditions of victory in chess are determined by rules, the *import* of victory, namely, what it *means* to win in chess, is not: this is not something we will find in a chess handbook, because we are assumed to know what it means to win a game. On the other hand, the constitutive rules of chess have to specify a set of conditions of victory, for otherwise we would not be able to understand in what sense chess is a competitive game. Hence, the meaning of the concept of victory depends not on the constitutive rules of chess but on the fact of chess being a competitive game. This is the sense in which concepts like victory can be called *meta*-institutional concepts: their relevance for the institution depends not on constitutive rules, as in the case of institutional concepts in the structural sense, but on the overall meaning of the system of rules as an instance of a given social practice.[4] Considering that this overall meaning shapes the system of rules within a teleological framework, I will call these kinds of meta-institutional concepts *teleological*. (And shortly I will be arguing that not all meta-institutional concepts are teleological.)

The teleological perspective is still internal in Hart's sense. It can be said to frame a more reasoned form of acceptance than the structural perspective, because it endows the system of constitutive rules with a clear meaning and rationale and so makes it possible to actually *understand* the game. But since this perspective is not just dependent on institutional concepts, a comprehensive, reasoned, well-grounded internal point of view toward an institution must be understood to include more than rule-constituted concepts: at a minimum, it needs to also include the meta-institutional concepts that make it possible to design a given system of rules in view of its objective. Hence if, as agents internal

[2] The distinction between an institution conceived as a system of constitutive rules and its broader meaning as a practice was first introduced by Hubert Schwyzer (1969). On this distinction see also Lorini 2000, 263ff., Marmor 2009, Roversi 2010. The teleological structure of rule-constituted institutional practices, and their connections with "basic needs of human life," has been stressed and discussed thoroughly by Wojciech Żełaniec in Żełaniec 2013, 106-108, 151-55. See also, in this regard, Ottonelli 2003.

[3] This has also been noted by Amedeo G. Conte (1995: 530), and subsequently by Giuseppe Lorini (2003: 299).

[4] The expression *meta-institutional*, along with its concept can be found in an insightful paper by Dolores Miller (1981) where she discusses the example of victory in competitive game-playing in analogy to the role the concept of obligation plays in Searle's (1969: 63ff.) theory of the "essential rules" of speech acts. Giuseppe Lorini discusses how meta-institutional concepts can be "a new category for social ontology" in Lorini 2014.

to a legal system, we were able to carry out acts and interact with objects constituted by legal rules but ignored the meaning of law as a practice—what the law is *for*, what purpose a legal system is meant to serve—we would still have an internal point of view, but one that corresponds to a sort of institutional autism, a self-referring attitude that would turn us into blindfolded technical legal machines. Such a purely structural perspective, detached from its teleological sense, is certainly not impossible: we can be motivated to apply the rules without pondering their purpose. But the self-reflective critical attitude that Hart speaks of in analysing the internal point of view can raise the question about the grounds we have for accepting the rule, and this will in turn require us to understand and accept—at least in outline—the rationale behind the rules. Hence, it seems safe to conclude that Hart's internal point of view on a given institutional structure can be grounded only if that structure is viewed through a combined structural and teleological perspective.

An example of a teleological meta-institutional legal concept that is particularly relevant for Hart's theory is that of legal validity. Legal systems typically lay down the main procedures by which to enact valid norms, but this validity is connected to specific institutional concepts denoting different kinds of directives.[5] In Italy, for example, there is a set procedure that Parliament needs to follow if it is to enact a *legge ordinaria* (a statute), but *legge* is a technical term in the Italian legal system, because there are other sources of law which are not strictly speaking *leggi*. Hence, this institutional framework can be reconstructed by defining *leggi* as kinds of directives whose institutional concept is created by constitutive rules: a text approved by both chambers of the Italian Parliament under a given procedure counts as a *legge* (status function) in the Italian legal system. Now, the import of a *legge* is to establish one or more legally valid norms, but this concept of validity is not constituted by any rules of the Italian legal system itself. Rather, it depends on the broader practice of legal norm-enactment: it is the typical outcome and main objective of norm-enactment in general within a parliament. Hence, while *legge* is an *institutional* concept, validity is a *meta*-institutional concept. And if we are to have an internal point of view on the institution of norm-enactment in the Italian legal system, we of course need to take *both* concepts into account.

4. The Axiological Perspective

Adriana likes chess very much, and she is very good at it. In fact she is *too* good. Given that her knights seem to be always in the right place, I pay closer attention to the way in which she moves them until I realize that she is moving them incorrectly: not by two squares and then a turn but by three or even four squares and then a turn. I correct her several times but she keeps moving her knights incorrectly. In the end I say, "Adriana, you can't move your knight to that square! That move is *illegal*". To which she replies, "But I *want* the knight to be on that square! It's almost checkmate that way!" I thus understand. She is not making mistakes. "You are cheating! You cannot cheat while playing: you're supposed to follow the rules!" "Why?" she asks naively. "You didn't mention this rule".

[5] I have developed this point (as well as several others related with the different perspectives) in a manuscript titled "Five Kinds of Perspectives on Legal Institutions", available online at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2143275

"That's true", I say. "And the reason I didn't is that I took it for granted, because to break it", I go on to explain haltingly, "would be, well, dishonest".

As in the case of the teleological perspective, I am here taking a point of view broader than the purely structural one. When I explained to Adriana that she was supposed to follow the rules in striving to win, I was making explicit something that matters not only in the game of chess but also in the wider practice of competitive game-playing: I was referring to the *background* of constitutive rules rather than to their structure. But here the background element I am referring to is not the point of the practice, or what is ostensibly its ultimate objective, but its axiological structure, or the complex of values that are generally associated with it. When playing a game, honest players follow the rules in trying to win: from an evaluative and moral point of view, it is far better to lose and play honestly than to cheat and win. This axiological perspective is connected with the teleological one, either because the objective of the practice can be understood as inherently valuable or because there is a specific deontology that must be observed in trying to achieve that objective. Where games are concerned, both perspectives are relevant: on the one hand, the objective of the game is to win, but on the other (in normal circumstances) it is to enjoy your time with your opponent. If you cheat, you are spoiling the enjoyment of your opponent in competing with you. The assumption here is therefore that the deontology embedded in game-playing is functional to the value of cooperation as an essential element making it possible to enjoy a shared activity or, in a strictly competitive setting, making it possible to compete on an equal footing.

When we adopt this axiological perspective, cheating present a concept relevant for chess. But, as in the case of victory, the concept of cheating is not constituted by the rules of chess. Rather, it is connected with certain features of the general game-playing practice. Moreover, while the constitutive rules of chess set out the conditions of victory, they do not set out any conditions subject to which a player may cheat, at least not directly.[6] Since cheating is not constituted by the rules of the game and depends on the general features of the game-playing practice, it is a *meta*-institutional concept. But it is an *axiological* meta-institutional concept and not a teleological one, because it is used in evaluative judgments that already take for granted which values the practice is meant to embody and which kinds of behaviour are consequently held to be offensive and morally objectionable.

Clearly, given the emphasis the axiological perspective lays on an institution's axiological structure, and hence on the values the institution is ultimately meant to serve, this perspective can figure as an aspect of the self-reflective critical attitude that Hart connects with the internal point of view. In a sense, this axiological perspective is even more internal than the teleological, because it inevitably brings evaluative judgments to bear, and in assessing the grounds for criticizing our own and others' behaviour, the axiology behind the institution will be even more important than its purpose. And no doubt this perspective is more internal than the purely technical structural perspective, because only by taking an axiological standpoint can the normative grounds of our acceptance

[6] To my knowledge, the best treatment so far made of the problematic relation between cheating and the constitutive rules of a game is that of Amedeo G. Conte: See Conte 2003.

ultimately be justified. This further supports the two claims previously made in discussing the teleological perspective, namely, that the internal point of view can be analytically broken down into constituent perspectives, and that judgments made from the internal point of view need more than institutional, rule-constituted concepts for their formulation. If we want to fully understand the internal point of view and its grounds, we will have to focus not only on secondary, constitutive rules but also on their conceptual and axiological import and background.

If we go back to the example of validity and norm-enacting procedures, we can clearly appreciate in what way axiological meta-institutional concepts and the axiological perspective are relevant for the internal point of view. In fact, apart from being "valid", statutes can also be distinctively just or unjust, and norm-enacting procedures can be evaluated in light of their capacity to realize substantive moral and political values such as democracy. It can even be argued that to a certain extent if a norm-enacting speech act openly contradicts its own claim to justice, it is defective if not self-defeating: this is Robert Alexy's (2002: 35ff.) well-known "argument from correctness", according to which an act of constitutional norm-enactment, like "X is a sovereign, federal, and unjust republic", is tantamount to a performative contradiction because it patently conflicts with the claim to correctness and justice implicitly made by all acts of legal norm-enactment (see Alexy 2002: 38). The problem, in this case, is not that the procedures for enacting a constitutional norm have not been followed correctly, but rather than something wrong happens *even if* we are following those procedures correctly. Hence, in order to work through this problem, or at least understand it, we need to conceive that act of norm-enactment as embedded in a general practice grounded in a specific set of values, and how these values are conceptually interconnected with the inner structure of that practice; that is, we have to view the law from an axiological perspective distinct from the purely structural one—distinct, to be sure, but even more internal.

One could observe here that the interpretation I am providing of Hart's internal point of view is cast in the mould of natural law theory, along the lines of the argument that John Finnis (1980: 13), for example, makes about the "central case" of the "legal viewpoint". I am, to be sure, assuming that to accept an institution's rules on the ground of their axiological and teleological rationale is to take a point of view which is more internal than simply accepting the rules such as they are, as exclusionary and entrenched reasons. I am not, however, assuming that an internal point of view is impossible outside a teleological or axiological perspective. Indeed, as noted, people can hold a purely *structural* internal point of view. What I am arguing, rather, is that a person holding a teleological and axiological perspective will have a deeper understanding of the grounds for the critical attitude connected with an internal point of view. Hence, rather than providing a natural-law interpretation of Hart's internal point of view and assuming that an appeal to the axiological dimension is "the central case of the legal viewpoint" (Finnis 1980: 15), I am trying to nuance the critical reflective attitude connected with the internal point of view: I am doing so by describing different degrees of internality in terms of a more or less deep acceptance of the

grounds for accepting an institution's rules—and that without assuming any cognitivist account of the axiological dimension behind an institution.[7]

5. The Strategic Perspective

Adriana is making progress. I see that she is now very cautious in moving her queen, whereas her previous game plan essentially consisted in capturing every piece by moving the queen all across the board. "I like the queen", she would say, "I want to capture pieces with my queen". But now she appreciates that it is better to avoid using the queen to capture pieces if there are other means of doing so. Not only has she correctly learned the rules of chess but she also uses them well in her game, because she understands the comparative advantage of the pieces when playing: she understands not only the structure of chess and its basic features as a game, but also the flow of a good game, appreciating, for example, when it is advisable to attack the opponent's king and when to defend her own.

Adriana is now able to have a strategic perspective on the game, enabling her to play it *well*, that is, to her own advantage in working on victory. This strategic perspective is not focused on the institution's structural features in the abstract but on actual practice: its basic question is how to maximize the likelihood of achieving the objective. As that suggests, the teleological aspect of the practice is crucial here, too, but the strategic perspective is entirely different from the teleological one, because on a strategic perspective we consider not only the main objective of the game and the structural conditions of its achievement but also what the most effective way is to achieve that objective in a concrete match.

The strategic perspective includes several specific concepts. Consider this entry on chess drawn from Wikipedia:

> The *King's Indian Attack* (KIA), also known as the *Barcza System* (after Gedeon Barcza), is a chess opening system for White, most notably used by Bobby Fischer. [...] The KIA is often used against the *semi-open defences* where Black responds asymmetrically to e4, such as in the *French Defence*, *Sicilian Defence*, or *Caro-Kann Defence*. Yet it can also be played against Black's more common *closed defenses*, usually through a move order that begins with 1. Nf3 and a later *fianchetto* of the white-square bishop. For this reason, transpositions to the *Réti Opening*, *Catalan Opening*, *English opening* or even the *Nimzo-Larsen Attack* (after b3 and Bb2) are not uncommon.[8]

Here we have a list of several terms referring to different kinds of attacks, defences, and openings in chess. Statements of this kind are quite normal in chess theory, but it is important to note that these concepts are not institutional. In fact, they are not constituted by the rules of chess but rather denote different ways in which the institutional, rule-constituted elements of chess can be used. These concepts are not meta-institutional, either, because they depend not on the features of the general game-playing practice but rather on the specific fea-

[7] I am grateful to Jaap Hage for pointing out to me this possible ambiguity of my conception.

[8] From the Wikipedia entry "king's Indian attack": http://en.wikipedia.org/wiki/King%27s_Indian_Attack. Italics added. Accessed January 4, 2017.

tures of chess: there is no other game where you can carry out a king's Indian attack. Concepts of this kind I will call *para-institutional*, using the prefix *para-* in the same sense as it is used in terms such as *paramedic*, *paralegal*, or *paralanguage*, namely, as qualifying objects which in some sense attach to more fundamental entities, and which are relevant for the concrete practice revolving around those entities. The same is true of para-*institutional* concepts: they work as parts of descriptive sentences that already take for granted the instantiation of one or more institutional elements. When, for example, I say that "Kasparov carried out a king's Indian attack", the truth conditions of this sentence include the truth of other descriptive sentences, such as "Kasparov moved a knight" and "Kasparov moved a pawn", and these other sentences are formulated from a structural perspective, in that they only involve institutional, rule-constituted concepts.

Now, is the strategic perspective internal or external in Hart's sense? I submit that it can be both. Of course we can adopt a strategic perspective from the internal point of view: players in any game do this ordinarily, but so do members of parliament, who usually act both in accordance with parliamentary procedure and with their concrete experience in maximizing the chances of having a legislative bill actually pass into law. But the strategic perspective can be external, too, because a "bad man" who should reject an institution's rules could still use them to advantage by exploiting the fact that almost everyone else accepts and follows the same rules. Typically, in this case, the point of view taken in such a scenario would be a moderately external one in Hart's sense, namely, a point of view from which "the observer [...], without accepting the rules himself, assert[s] that the group accepts the rules, and thus [...] from the outside refer[s] to the way in which *they* are concerned with them from the internal point of view" (Hart 1994, 89). In fact, in order to work out strategies by which to achieve an objective that is internal to an institution, a bad man will have to take the institution's internal structure and teleology into account, however much externally.

Let me now illustrate how a strategic perspective, with its connected para-institutional concepts, can be relevant for our previous example of norm-enactment in a parliament. Consider the case of parliamentary obstructionism, or "filibuster". The concept of filibuster denotes a set of parliamentary strategies that can be carried out in instantiating institutional rule-constituted concepts. In a filibuster, we follow all the procedures constituted by the rules of parliament, but we do so as a means of carrying out strategy, as when making attack or defence moves in chess. This, then, is a typical case of a para-institutional concept subordinate to the institutional concepts making up a legal system's norm-enactment procedures. When parliamentarians filibuster, they can do so from either an external or an internal point of view, but in either case their perspective will be strategic: they may be contrary to parliamentary practices in general, and hence act "from within" to take them down, or they may have constitutional reasons for rejecting a specific legislative bill, and hence act strategically to prevent it from passing.

6. The Sociological Perspective

Adriana and I have been playing chess for an entire year. She has become a very good player, probably better than me. She is fully conversant with all the rules and their variants, and she no longer cheats. In fact, chess has taught her the

value of tacit cooperative agreements to follow the same set of rules, even in competitive settings. I am very proud of her but cannot imagine just how deep her understanding of the game has become. One evening we are watching TV together when Adriana suddenly says, "Dad, chess is a good game, but it's not evenly balanced: the player who makes the first move always has a slight advantage". "No, that's an impression and it's mistaken", I reply: "Chess is the perfect game. It's common knowledge". But then I check Wikipedia, and it supports her statement:

> The first-move advantage in chess is the inherent advantage of the player (White) who makes the first move in chess. Chess players and theorists generally agree that White begins the game with some advantage. Since 1851, compiled statistics support this view; White consistently wins slightly more often than Black, usually scoring between 52 and 56 percent.[9]

In considering whether the first player has a slight advantage in a chess match, I am not looking at any rule of the game or any combination of rules, and of course I am not discussing the typical teleology or values of competitive gameplaying. Nor am I considering a strategy, because the question here is theoretical rather than practical: it is not a matter of how to maximize my chances of winning but a matter of how chess matches concretely unfold in general. This is a sociological perspective on the institution, where we focus on the patterns and features of an actual institutional practice that can be generalized on a statistical basis: we are looking at features of an institution as an actual practice over against its structural features as constituted by its own rules.

A sociological perspective generates its own concepts. The concept of first-move advantage is not constituted by the rules of chess itself but rests on the way in which chess winds up looking like when actually played. Just like the concept of "king's Indian attack", that of first-move advantage *depends* on constitutive rules but is not itself *constituted* by those rules: it needs the rules of chess to define the game's constitutive elements, for otherwise it could not have the meaning it has, but its features are owed to the way in which the game of chess unfolds in actual gameplay. Thus, the concept of "first-move advantage" is a para-institutional concept, but a theoretical and not a practical one, because it denotes something which typically *happens* and not anything that one can *do*.

The sociological perspective and its associated para-institutional concepts can be described more perspicuously by enlisting the help of Searle's theory. In *Making The Social* World, Searle treats the problem of "systematic fallouts", underscoring several crucial properties of these phenomena, the most important being that, while institutional phenomena depend (in his theory) on collective acceptance, systematic fallouts do not: they are "intentionality-independent facts about intentionality-relative phenomena" (Searle 2010: 117).[10] Further, Searle notes that, while institutional elements typically have "deontic powers" or normative consequences, systematic fallouts do not have such consequences. Thus, systematic fallouts depend on "ground-floor institutional facts" and carry no distinctive deontology, to the point that participants in a given institutional practice

[9] The Wikipedia entry "first-move advantage in chess: https://en.wikipedia.org/wiki/First-move_advantage_in_chess. Accessed January 4, 2017.
[10] See also, in this regard, Thomasson 2003: 275-56; Andersson 2007: 105-26.

can very well be unaware of them. Searle makes in this regard a baseball example very similar to our example of the first-move advantage:

> To take a trivial example, it has been discovered in baseball that, statistically, left-handed batters do better against right-handed pitchers, and right-handed batters do better against left-handed pitchers. This is not required by the rules of baseball; it is just something that happens. I propose to call these "third-personal fallout facts from institutional facts," or more briefly, "fallouts" from institutional facts. They are "third-personal," because they need not be known by participants in the institution. They can be stated from a third-person, anthropological, point of view. They carry no additional deontology, and so no new power relations are created by fallouts (Searle 2010: 117).

As Searle notes, systematic fallouts are typically described in economic theory:

> In economics the ground-floor facts are in general intentionality-relative. For example, so and so bought and sold such goods. But the facts reported by economists are typically intentionality-independent. For example, the Great Depression began in 1929 (Searle 2010: 117).

Just like the strategic perspective, the sociological perspective from which we consider systematic fallouts denoted by way of theoretical para-institutional concepts can be either internal or external. For example, a teacher of constitutional law can accept the overall constitutional system but need to also consider and describe the kinds of features and possible flaws the corresponding institutional framework entails in actual practice. On the other hand, an alien observer can take a sociological perspective from an external point of view, to the point of considering in statistical terms not only the distinctive "fallouts" of an institutional practice but also normal rule-abiding behaviour. This would be an extreme external point of view in Hart's sense, in which "an institutional practice is described in terms of observable regularities of conduct, predictions, probabilities, and signs" (Hart 1994: 89-90). Finally, it is perfectly possible to hold a sociological perspective from a moderate external point of view. Indeed, any bad man willing to act in an institutional framework strategically, thus considering the most effective way of achieving an objective in light of the beliefs that people hold, will be exploiting observations made from a sociological perspective. Hence, even if observers typically take a sociological perspective, whereas participants take a strategic one, participants who act strategically can advance their strategy by reasoning from a sociological perspective. On the other hand, a good sociological description of how a given legal framework works in practice needs to be able to consider the strategic reasoning of all actors. Thus, the two perspectives that typically involve para-institutional concepts—the strategic and the sociological—are intertwined in a peculiar way, as are the teleological and axiological perspectives, which involve meta-institutional concepts.

The relevance of theoretical para-institutional concepts can be illustrated by going back once more to our example of legal norm-enactment. Some situations, such as filibuster, can give rise to "legislative gridlock", namely, situations where actual parliamentary practice cannot have any significant normative outcome because no party has a filibuster-proof majority. This is a para-institutional concept, because the constitutive rules of parliamentary procedure do not *create* the concept themselves but are *necessary* to create the institution in which legisla-

tive gridlock can happen. However, unlike filibuster, which is a practical concept relevant to the participants engaged in the practice, legislative gridlock denotes a typical situation created by parliamentary practice, not a kind of parliamentary strategy. Through this theoretical para-institutional concept we are describing the possible outcomes of an institutional practice from a sociological perspective. Of course, as noted, we can momentarily take this perspective for strategic purposes: if for, example, deputies know how legislative gridlocks can happen in a given constitutional framework, they can use this information to act strategically toward the goal of forcing new elections.

7. Constitutive Rules between Background and Practice

The description so far made of the various perspectives on an institution makes it possible to analyse institutional ontology in a more fine-grained way than can be achieved on the basis of the simple model of constitutive rules. In fact, constitutive rules are the crucial element of institutions only from a structural perspective: if we consider institutions from the other perspectives previously described, we will see that there are at least two crucial elements that need to be taken into account. The first of these is how constitutive rules are related to their *background*; the second how these rules translate into *practice*.

On the one hand, the teleological and axiological perspectives show that constitutive rules can create a set of institutional concepts—and hence an institutional activity (chess, for example)—only against the background of another, more fundamental kind of activity (competitive game-playing, in the case of chess): they frame a surface layer that must be embedded within a deeper layer.[11] This background dictates specific conceptual and normative boundaries within which constitutive rules can be framed. For example, the constitutive rules of a competitive game *must* define conditions of victory and normally presuppose that players will not cheat. Hence, the teleological and axiological perspectives make it possible to ascertain several considerations that have an *a priori* impact on the institutional structure, because they define the conceptual and normative background within which constitutive rules operate.

On the other hand, the strategic ad sociological perspectives show that there is a difference between the structure created through rules, namely, the institution on paper, and the way in which constitutive rules are actually put into practice, namely, the institutional activity in its concrete unfolding. In a sense, and somewhat mysteriously, there always develop emergent features of an institutional practice that cannot be foreseen simply by looking at the institutional structure itself. These emergent features can in the long run entail structural considerations. For example, we could seek to change the structural features of a given norm-enacting process if we find that some strategy can be abused, or phenomena considered from a sociological perspective can eventually result in the death of that institution and hence require a massive overhaul of its constitutive rules. Hence, just as in the case of considerations made from the axiological and teleological perspectives, even those made from a strategic or sociological perspective can affect the institutional structure. But, unlike the former, they do

[11] I am borrowing from Marmor 2009 this distinction between surface and deep conventions.

so only *a posteriori*, that is, only after the institution has had some history of practice.

Institutions therefore have a distinctive three-dimensional ontology: First, they are framed against an *a priori* conceptual and normative *background*; second, their constitutive rules define a *structure*; and third, they have *a posteriori* emergent features resulting from their *practical outcomes*, or from what happens when they are put into practice. Constitutive rules provide only one of these three elements, and for an adequate explanation of institutions, they must therefore be complemented by the other two.[12]

This applies not only to constitutive rules but also to Hart's idea of the internal point view. In our analysis of the three perspectives from which an institution may be viewed, we have seen how these perspectives and the three dimensions of institutional ontology interlock with Hart's dichotomy between the internal and the external point of view (and with his further dichotomy between the moderate and the extreme external point of view). The analysis essentially brings out the further shades that can be brought to the main idea: Hart's threefold distinction between an internal, a moderate external, and an extreme external point of view becomes a distinction between six possible approaches and corresponding characters—two for each of the three points of view, depending on the kinds of perspectives that are dominant for that character.

The internal point of view can be shaded into the perspectives of two characters: the *committed practitioner* and the *alienated technician*. The main difference between the two is that, while the former has a structural, teleological, and axiological perspective on institutions, the latter looks at them solely through the structural perspective. A committed practitioner takes a deep internal point of view on an institution and considers both its structure and its axiological and conceptual background, not only accepting the rules but also understanding why they have been framed that way, in light of their basic conceptual boundaries and values. An alienated technician takes a superficial internal point of view on the institution and considers only its structure, without necessarily considering in full how this structure is connected with deeper requirements and values. This person knows the institution's constitutive rules very well and follows them closely, but without taking their conceptual background into account: constitutive rules are understood by this character as reasons for action in themselves, without having to take up the question of their justification, and hence without taking a fully self-reflective attitude to their normative force.

The moderate external point of view can likewise be shaded into the perspectives of two characters: the *bad man* and the *social theorist*. Here the main difference is that, while the former acts primarily from a strategic perspective, the latter instead acts from a sociological perspective. Both the bad man and the social theorist take account of the way people behave within a certain institutional framework in light of their beliefs about that institution. They are interested in what committed people believe about the institution and so can also take a structural perspective, or even a teleological and axiological one. But, for them,

[12] That law has a three-dimensional ontology has been pointed out before by several legal philosophers in the 20th century, three classic examples being Gustav Radbruch (1993), Hermann Kantorowicz (1962: 69-70), and Miguel Reale (1968), and it would be interesting to see how these conceptions relate to the analysis of institutions offered in this paper.

any consideration made from the structural, teleological, or axiological perspective is subordinate to the strategic or the sociological one: both the bad man and the social theorist are interested in what people believe about the institution, but while the former needs this information to maximise the likelihood of achieving the desired institutional outcome, the latter needs it in order to understand how such beliefs translate into concrete social phenomena. Constitutive rules, and their relevant conceptual and axiological background, do not provide them with normative reasons for action: the bad man considers them as providing prudential reasons for action, while the social theorist simply describes them.[13]

Finally, even the extreme external point of view can be shaded into the perspectives of two characters: the *alien agent* and the *alien scientist*. In one respect, the difference between them can be analogised to the one between the bad man and the social theorist, in that, while the alien *agent* acts from a strategic perspective, the alien *scientist* reasons from a sociological one. But in another respect the analogy breaks down, because these two perspectives afford an entirely different view once we step into the shoes of the latter two characters. Which is to say that, unlike the former pair, the alien agent and the alien scientist both confine themselves to recording the regularities of behaviour that people have within a certain institutional framework and assign to them a certain degree of statistical probability: they do so without entering into the conceptual domain of that institution. Hence, in this case, no reference is made to the structural, teleological, or axiological perspectives. For the alien agent, considerations made from the strategic perspective can work as prudential reasons for action, whereas for the alien scientist considerations made from the sociological perspective are simply descriptions and do not involve any kind of practical reasoning.

In this way, by appealing to the different perspectives previously identified, Hart's threefold distinction between points of view on an institution is further shaded into a sixfold distinction. These six perspectives differ by reason of the way each of the characters who embody them sees the constitutive rules of a given institution: the committed participant considers them as strongly normative reasons for action, because he or she understands and embraces their conceptual and axiological background, whereas the alienated technician considers them as normative reasons for action in isolation, namely, simply as practical assumptions; the bad man considers them as prudential reasons for action, whereas the social theorist considers them as the content of social descriptions; finally, neither the alien agent nor the alien scientist considers them at all.

8. Closing Remarks

In this paper, I proposed that Hart's concepts of the internal point of view and of secondary rules be combined with Searle's concept of constitutive rules. As discussed, this promises to be a powerful paradigm for analysing institutional on-

[13] Of course this reference to a "bad man's" attitude goes back to Oliver Wendell Holmes (1897: 459). Stephen Perry (2000: 164) has argued that Holmes's bad man adopts not an external point of view but an internal one, because he is engaged in practical and not theoretical reasoning, a sort of practical reasoning based on prudential reasons. It seems to me that prudential reasons cannot ground an internal point of view in Hart's sense, and that an external point of view can include both a practical (strategic) and a theoretical (sociological) attitude.

tology. What I have argued is that this paradigm can be further enriched and specified. Constitutive rules create a structure embedded within a conceptual and axiological background and give place to a concrete activity that always ends up having emergent features: hence, they are connected both with a background and with a practice. Background, structure, and practice are thus the core elements of a three-dimensional institutional ontology. I have argued for this conclusion by distinguishing between four kinds of perspectives on an institution, and this in turn made it possible to further specify (or shade) Hart's threefold distinction between an internal, a moderate external, and an extreme external point of view into a sixfold distinction between the perspectives taken by possible paradigmatic characters who represent different ways of approaching the institution in question.

References

Alexy, R. 2002, *The Argument from Injustice. A Reply to Legal Positivism*, Oxford: Clarendon Press.

Andersson, A. 2007, *Power and Social Ontology*, Malmö: Bokbox Publications.

Conte, A.G. 1995, "Deontica wittgensteiniana", in A.G. Conte, *Filosofia del linguaggio normativo. II. Studi 1982-1994*, Torino: Giappichelli, 517-61 (1st ed. 1993).

Conte, A. G. 2003, "Filosofia del baro", *Rivista internazionale di filosofia del diritto*, 80, 679-747.

Finnis, J. 1980, *Natural Law and Natural Rights*, Oxford: Clarendon.

Hart, H.L.A. 1994, *The Concept of Law*, 2nd edition, Oxford: Clarendon (1st ed. 1961).

Holmes, O.W. 1897, "The Path of the Law", *Harvard Law Review*, 10, 457-78.

Kantorowicz, H.U. 1962, "Staatsauffassungen (Conceptions of the state)", in T. Würtenberger (ed.), *Rechtswissenschaft und Soziologie: Ausgewählte Schriften*, Karlsruhe: Müller, 69-81 (1st ed. 1925).

Lorini, G. 2000, *Dimensioni giuridiche dell'istituzionale*, Padova: Cedam.

Lorini, G. 2003, "Cinque livelli di descrizione di un atto istituzionale", in Di Lucia, P. (ed.), *Ontologia sociale. Potere deontico e regole costitutive*, Macerata: Quodlibet, 299-308.

Lorini, G. 2014, "Meta-institutional Concepts. A New Category for Social Ontology", *Rivista di Estetica*, 54, 127-39.

Marmor, A. 2009, *Social Conventions: From Language to Law*, Princeton: Princeton University Press.

Miller, D. 1981, "Constitutive Rules and Essential Rules", *Philosophical Studies*, 39, 183-97.

Ottonelli, V. 2003, "Funzioni di 'status', linguaggio e regole costitutive. Potere deontico e regole costitutive", in Di Lucia, P. (ed.), *Ontologia sociale. Potere deontico e regole costitutive*, Macerata: Quodlibet, 309-31.

Perry, S. 2000, "Holmes versus Hart: The Bad Man in Legal Theory", in S.J. Burton (ed.), *The Path of the Law and Its Influence. The Legacy of Oliver Wendell Holmes, Jr.*, Cambridge: Cambridge University Press, 158-96.

Radbruch, G. 1993, "Grundzüge der Rechtsphilosophie (Outlines of legal philosophy)", in *Gesamtausgabe*, *Vol. 2*. Heidelberg: Müller, 9-204.

Reale, M. 1968, *Teoria Tridimensional do Direito*. São Paulo: Saraiva.

Roversi, C. 2010, "Constitutive Rules in Context", *Archiv für Rechts und Sozialphilosophie*, 96, 223-38.

Schwyzer, H. 1969, "Rules and Practices", *The Philosophical Review*, 78, 451-67.

Searle, J.R. 1969, *Speech Acts*, Cambridge: Cambridge University Press.

Searle, J.R. 2010, *Making the Social World: The Structure of Human Civilization*, Oxford: Oxford University Press.

Thomasson, A. 2003, "Foundations for a Social Ontology", *Protosociology*, 18-19, 269-90.

Żełaniec, W. 2013, *Create to Rule: Studies on Constitutive Rules*, Milan: LED.

Advisory Board

SIFA former Presidents

Eugenio Lecaldano (Roma Uno University), Paolo Parrini (University of Firenze), Diego Marconi (University of Torino), Rosaria Egidi (Roma Tre University), Eva Picardi (University of Bologna), Carlo Penco (University of Genova), Michele Di Francesco (IUSS), Andrea Bottani (University of Bergamo), Pierdaniele Giaretta (University of Padova), Mario De Caro (Roma Tre University), Simone Gozzano (University of L'Aquila), Carla Bagnoli (University of Modena and Reggio Emilia), Elisabetta Galeotti (University of Piemonte Orientale)

SIFA charter members

Luigi Ferrajoli (Roma Tre University), Paolo Leonardi (University of Bologna), Marco Santambrogio (University of Parma), Vittorio Villa (University of Palermo), Gaetano Carcaterra (Roma Uno University)

Robert Audi (University of Notre Dame), Michael Beaney (University of York), Akeel Bilgrami (Columbia University), Manuel Garcia Carpintero (University of Barcelona), José Diez (University of Barcelona), Pascal Engel (EHESS Paris and University of Geneva), Susan Feagin (Temple University), Pieranna Garavaso (University of Minnesota, Morris), Christopher Hill (Brown University), Carl Hoefer (University of Barcelona), Paul Horwich (New York University), Christopher Hughes (King's College London), Pierre Jacob (Institut Jean Nicod), Kevin Mulligan (University of Genève), Gabriella Pigozzi (Université Paris-Dauphine), Stefano Predelli (University of Nottingham), François Recanati (Institut Jean Nicod), Connie Rosati (University of Arizona), Sarah Sawyer (University of Sussex), Frederick Schauer (University of Virginia), Mark Textor (King's College London), Achille Varzi (Columbia University), Wojciech Żełaniec (University of Gdańsk)

Presentism and Causal Processes

Ernesto Graziani
University of Macerata

Abstract

Presentism is the view that only present temporal entities (tenselessly) exist. A widely-discussed problem for presentism concerns causation and, more specifically, the supposed cross-temporally relational character of it. I think that the best reply to this problem can already be found in the literature on temporal ontology: it consists, roughly, in showing that (at least) some of the main approaches to causation can be rephrased so as to avoid commitment to any cross-temporal relation, including the causal relation itself. The main purpose of this paper is to extend this reply to the process view, an approach to causation that has not been considered within this debate until now. I shall do this by taking into account Dowe's conserved quantity theory—a recent and prominent theory of this sort—and employing it as a proxy for the other major process theories of causation. In dealing with Dowe's process theory of causation, however, two additional problems must be faced: one concerns the four-dimensional spacetime framework on which its formulation relies; the other concerns the very notion of causal process (and the companion notion of causal interaction). While the presentistic account of Dowe's theory (and, virtually, of the process view of causation in general) put forth in this paper is intended primarily as a contribution to the mentioned paraphrase-based enterprise of reconciliation between presentism and causation, I shall also offer some reasons for presentists to prefer the process view of causation to the other views of causation that have already been reconciled with presentism.

Keywords: Causal processes, Causation, Cross-temporal relations, Presentism.

1. Introduction

Presentism is the metaphysical view that only present temporal entities (tenselessly) exist. (Presentism contrasts with a variety of views on time; for the purpose of this paper, however, it suffices to mention just the main opponent of presentism: *eternalism*, the view that past and future entities (tenselessly) exist as well.)[1] A widely-discussed problem for presentism regards causation—an aspect of reality playing a pivotal role in many branches of science and philosophy, and in ordinary thought as well. Part of this problem finds expression in the *argu-*

[1] As to the current "triviality debate", in line with Sider (2001), Hestevold and Carter (2002), Torrengo (2012), and others, I think that the predicate 'exists' occurring in the definition of presentism (and of the other theories in temporal ontology) should be read as tenseless, i.e., as expressing an attribution of existence deprived of tense.

Argumenta 4, 1 (2018): 159-176
ISSN 2465-2334
© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.gra

ment from causation, which runs as follows.[2] Causation is a relation taking events as relata. In many—if not all—cases, the causal relation is *cross-temporally exemplified* or, more briefly, *cross-temporal*, i.e., exemplified by its relata at different times. So, in many—if not all—cases, if one of the causal relata exemplifies the causal relation at the present time, the other one exemplifies it at a past time or at a future time.[3] But for a relation to hold between two (or more) entities, it is required that the entities it relates (tenselessly) exist (at the times they exemplify the relation). Hence, presentism must be false. As anticipated, however, the problem that presentism has with causation is captured by the argument from causation only partially. In fact, depending on the various ways of detailing its nature, causation may be thought to involve various further relations (e.g., precedence, temporal contiguity, spatial contiguity, momentum transfer) between the cause-event and the effect-event or between the constituents of them (e.g., objects and times). Of course, if any of these further relations prove to be cross-temporal—and some of them surely do so, e.g., the precedence relation, which is inherently cross-temporal—then they result to be problematic for presentism just like the causal relation.[4]

Various philosophers—Sider (1999), Crisp (2005), Bourne (2006: 109-15), Brogaard (2006) and (2013), and McDaniel (2010)—have addressed this problem by showing that at least some of the main views about causation—the regularity view, the counterfactual view, and the primitivistic view have been considered—can be reformulated so as to avoid commitment to any cross-temporal relation, not even to the causal one.[5] In this paper, the sort of reply put forward by these authors will be extended to a fourth major view of causation: the *process view of causation*.[6] This will be done by taking into consideration the version of it elaborated by Dowe (2000), namely the *conserved quantity theory* or *CQ theory*, and employing it as a proxy for any of the major process theories of causation. The idea is that the presentistic reformulation of Dowe's CQ theory expounded in this paper may be re-employed, with some appropriate modifications, to presentistically account for the other major process theories of causation as well. This way of proceeding seems licit for two complementary reasons: first, the CQ theory is one of the most recent and accurately worked-out process theories, and thus a good specimen of this kind of approach to causation; second, the CQ theory shares important structural similarities with the other major process theories: especially with recent proposals such as those of Salmon (1997 and 1998: 13-24)

[2] Other versions of this anti-presentistic argument are offered by Bigelow (1996), Crisp (2005), Bourne (2006: 109, 110), and McDaniel (2010).

[3] I write 'in *many*—if not all—cases' in order not to exclude in principle possible cases of simultaneous causation, i.e., cases in which the causal relation is exemplified by its relata at the same time.

[4] In this connection, it is worth mentioning that the argument from causation is part of the broader *argument from cross-temporal relations*, in which any cross-temporal relation is put forward as troublesome for presentism (hence, for example, besides causal relations, intentional, resemblance, semantic, and precedence relations).

[5] To be exact, the regularity view has been addressed by Sider (1999), Crisp (2005), and Bourne (2006: 110-13); the counterfactual view by Crisp (2005), Bourne (2006: 113-14), Brogaard (2006), and McDaniel (2010); the primitivistic view by Sider (1999), Crisp (2005), Bourne (2006: 114-15), Brogaard (2006) and (2013), and McDaniel (2010).

[6] I am not a presentist; in this paper, however, I put myself in the shoes of the presentist and defend presentism.

and Kistler (2006), but also—although to a lesser extent—with less recent proposals such as those of Aronson (1971), Fair (1979), and Castañeda (1980).

In recasting Dowe's process theory of causation according to a presentistic perspective, however, two problems must be addressed in addition to the one regarding cross-temporal relations. One concerns the fact that Dowe's CQ theory is originally formulated within the theoretical framework of Minkowski spacetime, which is a notoriously hostile environment for presentism. The other additional problem concerns the central notion of causal process and the companion notion of causal interaction, both of which designate sorts of entities that cannot fit in their entirety into the instantaneous present of the presentistic universe. As will be shown, this latter problem is an instance of a wider problem that presentism has with *durative events*, i.e., events taking a nonzero amount of time to occur.

This paper is organised as follows: §2 outlines the core content of Dowe's theory of causation (making a few little terminological changes and theoretical adjustments) and mentions some merits of the process view of causation both in general as a theory of causation and specifically for the presentist;[7] §3 addresses the problem of the Minkowskian spacetime framework; §4 addresses the problem of causal processes, causal interactions, and durative events *qua* temporally extended entities; §5 addresses the problem of cross-temporal relations; §6 concludes by offering a presentistic reformulation of Dowe's analysis of the grounds of causation and by laying bare the main controversial assumption underlying the solution adopted.

2. Essentials of Dowe's Conserved Quantity Theory of Causation

The key idea of the process view of causation is that a causal relation between two events must be accounted for by resorting to the causal processes and the causal interactions linking them. A *causal process* is the possession of a causally relevant physical property by an object through space and time (or spacetime) or the transfer of such a property by means of an object through space and time (or spacetime). A *causal interaction* is a spatial and temporal (or spatiotemporal) overlapping of two or more causal processes that involves the exchange of a causally relevant physical property between the constitutive objects of them or the transfer of such a property from the constitutive object of one causal process to the constitutive object of another. These notions may be taken to form the basic conceptual structure that, beyond differences in development and terminology, is common to the various process theories of causation.

In Dowe's CQ theory, the notions of causal process and causal interaction are defined as follows (2000: 90; for a reason that will be adduced in a moment, some amendments are made):

> CQ1. A *causal process* is a world line [more exactly: a world tube] of an object possessing a conserved quantity.
> CQ2. A *causal interaction* is an intersection of world lines [world tubes] that involves exchange of a conserved quantity.

[7] The exposition of Dowe's theory will be inevitably concise, focused on those elements having a relevance to the issue of its compatibility with presentism. The reader new to this theory is advised to refer directly to Dowe's works, especially to Dowe 2000 (chapters 5-8).

The notions of process and world line are defined as follows (2000: 90, 91; again, with some amendments):

> A *process* is the world line [world tube] of an object, regardless of whether or not that object possesses conserved quantities. […] A *world line* [just like a *world tube*] is the collection of points on a spacetime (Minkowski) diagram that represents the history of an object. This means that processes are represented by elongated regions, or 'worms', in spacetime.

The difference between a world line and a world tube is the following: a world line is the spatiotemporal path of an ideal point-like particle (not a real particle); a world tube is the spatiotemporal path of a real spatially extended object. And here is the announced reason to amend Dowe's exposition: the objects involved into the CQ theory belong to the latter kind, not to the former, as Dowe's very definition of object makes clear (2000: 91):

> An *object* is anything found in the ontology of science (such as particles, waves and fields), or common sense (such as chairs, buildings and people).

For this reason, talk of world lines should be replaced by talk of world tubes throughout the exposition of the CQ theory. The notions of conserved quantity, intersection, and possession, which are also employed in CQ1 and CQ2, are defined as follows (2000: 91, 92):

> A *conserved quantity* is any quantity that is governed by a conservation law, and current scientific theory is our best guide as to what these are. For example, we have good reason to believe that mass-energy, linear momentum, and charge are conserved quantities […]
>
> An *intersection* is simply the overlapping in spacetime of two or more processes. The intersection occurs at the location consisting of all the spacetime points that are common to both (or all) processes. An *exchange* occurs when at least one incoming, and at least one outgoing process undergoes a change in the value of the conserved quantity, where 'outgoing' and 'incoming' are delineated on the spacetime diagram by the forward and backward light cones, but are essentially interchangeable. The exchange is governed by the conservation law, which guarantees that it is a genuine causal interaction. It follows that an interaction can be of the form *X, Y, λ,* or of a more complicated form.
>
> *'Possesses'* is to be understood in the sense of 'instantiates'. An object possessing a conserved quantity is an instance of a particular instantiating of a property. We suppose that an object possesses energy if science attributes that quantity to that body. It does not matter whether that process transmits the quantity or not, nor whether the object keeps a constant amount of the quantity.

According to Dowe, the relata of the causal relation are *states of affairs* conceived of along the lines of Armstrong (1997), i.e., as exemplifications of attributes (properties or relations) by particular objects (Dowe 2000: 168, 169). States of affairs may be either facts or events (2000: 169,170):

> An *event* is a change in a property of an object at a time, for example, a quantitative change; or a related simultaneous change in more than one property of more than one object at a time, and so on. […] A *fact* is an object having a property at

> a time or over a time period. Because both events and facts concern objects, this fits well with the Conserved Quantity Theory.

I shall adopt, however, a partly different terminology, which I find more suitable and organic. Where Dowe talks of states of affairs, facts, and events, I prefer to talk, respectively, of *events*, *static events* and *dynamic events* (in line with Casati and Varzi 2006). This is a merely terminological departure from Dowe for the two typologies perfectly match with each other. In the rest of this section, Dowe's original terminology will be flanked to the one I favour; in the following sections of the paper, replaced by it.

For the purposes of this paper, it is important to establish a further distinction between kinds of events. As a time may be either an instant (i.e., a time of zero duration) or a period (i.e., a time of nonzero duration), we can distinguish between *instantaneous* events and *durative* events. Moreover, durative events can be plausibly considered as mereological sums of shorter and shorter events and, ultimately, of (infinite) instantaneous events ordered in temporal sequence (just like a period of time can be considered as composed of shorter and shorter periods and, by the end, by infinite instants in sequence).[8] As Dowe remarks, the attribute-exemplification view of events (states of affairs, in his terminology) "fits well with the Conserved Quantity Theory"; it may be added that causal processes and causal interactions may in fact be considered as durative events of particular sorts: causal processes as those consisting in the possession of a conserved quantity by a physical object; causal interactions as those consisting in the exchange of a conserved quantity by two or more physical objects.

Dowe's view also includes a form of physicalism, and it must be so if the CQ theory is to be considered as a conception of causation, not simply of physical causation. As Dowe writes (2000: 170):

> [S]uch facts [static events] or events [dynamic events], if they enter into causation, must involve conserved quantities or supervene on facts and events involving conserved quantities. For example, the fact that the ball is green must supervene on the fact that various bits of the surface of the ball have certain physical properties by virtue of which the ball looks green. If these properties are not conserved quantities, then they in turn must supervene on conserved quantities. This seems to be a natural development of the Conserved Quantity theory.

Albeit natural, this physicalistic development is not compulsory for those interested in Dowe's CQ theory, which could be accepted as an analysis of physical causation, not of causation *tout court*.

To represent a static event (fact, in the original terminology) consisting in the possession of a specific amount of a quantity, Dowe introduces formulae like '$q(a) = x$' or, more briefly, '$q(a)$', reading 'object a has (x amount of) conserved quantity q'; if a second conserved quantity is involved, it is expressed by 'q''' (2000: 170). To represent a dynamic event (event) consisting in the change in amount of a quantity, Dowe resorts to formulae such as '$\Delta q(a)$', which presuma-

[8] The two distinctions intersect: since a change—in the simplest case—consists in the exemplification of two incompatible properties by the same object at two different instants, instantaneous events cannot be dynamic but only static, while durative events may be either dynamic or static.

bly reads 'object a undergoes a variation in the amount of quantity q' (Dowe is not completely perspicuous on this point). To express a causal interaction, e.g., one involving objects a and b, and the quantity q, Dowe simply writes "the interaction $\Delta q(a)$, $\Delta q(b)$".

In defining the grounds of the causal relation—i.e., in giving necessary and sufficient conditions for the causal relation to hold—Dowe, for simplicity, only takes into account the case where the causal relata are static events (facts, in his original terminology).[9] Where 'a' and 'b' represent two objects, and 'q' and 'q''' represent two conserved quantities, the causal relation is analysed as follows:

> There is a causal connection (or thread) between a fact [a static event] $q(a)$ and a fact [a static event] $q'(b)$ if and only if there is a set of causal processes and interactions between $q(a)$ and $q'(b)$ such that:
> (1) any change of object from a to b and any change of conserved quantity from q to q' occur at a causal interaction involving the following changes: $\Delta q(a)$, $\Delta q(b)$, $\Delta q'(a)$, $\Delta q'(b)$; and
> (2) for any exchange in (1) involving more than one conserved quantity, the changes in quantities are governed by a single law of nature (Dowe 2000: 171, 172).

Although this analysis of the grounds of the causal relation involves exactly two objects and two quantities, it can be very easily adapted to cases involving any number of objects and quantities. On the one hand, it can be adapted to cases where only one object is involved by setting $a = b$ (in which case there would be no causal interaction) or only one quantity by setting $q = q'$. On the other hand, it can be adapted to cases where more than two objects or quantities are involved by modifying condition (1), and precisely by adding, for any further quantity or object, the corresponding changes in that quantity possessed by that object. Condition (2) is introduced by Dowe in order to exclude cases where more independent causal interactions occur accidentally in the same place and time (e.g., the case where two billiard balls collide and at that very moment one of the two emits an alpha particle: in this case the emission of the alpha particle is not the effect of the collision because emission of alpha particles and collision between macroscopic objects are governed by different laws of nature).

It should be noticed that neither condition (1) nor condition (2) indicates, of $q(a)$ and $q'(b)$, which one is the cause and which one is the effect. As Dowe

[9] Dowe does not worry to identify a single precise form that causal claims should take (in general or specifically in the CQ theory); in fact, he resorts to a variety of different formulations such as 'there is a causal connection between a fact $q(a)$ and a fact $q'(b)$', '$q(a)$ and $q'(b)$ are linked by a causal connection', 'the quantity of a is causally responsible for the quantity of b' (see 2000: 171-73). If, in looking for a single precise formulation, we stick to Dowe's explanation of locutions like '$q(a) = x$' as *sentences*, we face a syntactical problem: the predicate 'causes' (just like 'is causally connected to' and the like) cannot take sentences as arguments (it would be a syntactical mismatch) but only singular terms. We may fix this problem very easily by introducing an "adapter" symbol that operates turning sentences into singular terms. For example, we may convene that, where P is a sentence expressing some event, $[\![P]\!]$ is the event expressed by P; in other terms, '$[\![P]\!]$' reads as 'the event expressed by 'P''. This is only a minor adjustment, which does not add anything substantial to Dowe's account. It allows, however, to formulate causal claims in a more precise fashion. We may write, e.g., '$[\![q(a) = x]\!]$ causes $[\![q'(b) = y]\!]$' and read it 'the event expressed by '$q(a) = x$' causes the event expressed by '$q'(b) = y$''.

(2000: 110) tells us, the CQ theory is symmetric with respect to time and thus noncommittal on the issue of causation's direction, being purposely designed to allow for backwards causation (the reality of which seems to be suggested by quantum mechanics). To account for the prevailing direction of causation in time (i.e., for the fact that, typically, causes are earlier than their effects), the CQ theory must be thus supplemented in some way, and the way chosen by Dowe's is to resort to a version of Reichenbach's fork asymmetry (2000: chapter 8).

Before proceeding with my presentistic re-elaboration of the CQ theory, I must spend a few words to explain why reconciling the process view of causation with presentism is something worth carrying out despite the fact that various views of causation have already been reconciled with it.[10] Since the process view of causation represents one of the standard approaches to causation (along with the other views that have already been reconciled with presentism),[11] a presentistically suitable interpretation of it would constitute an appreciable contribution to the global project of reconciling presentism and causation. This, in my opinion, is the major reason to engage in a presentistic account of the process view.

But the process view of causation has some merits that, perhaps, make it even preferable, both in general *qua* theory of causation and specifically *for a presentist*, to the three views of causation that have already been reconciled with presentism. Here are some of them. First, the idea that it is a causal process (a causal "thread" or "rope"), not simply a causal chain of discrete events, that leads from the cause to the effect is very intuitive, also from a common-sense perspective. Second, a physics-based interpretation of causation, while probably extraneous to common-sense, will be appealing to naturalistically inclined philosophers and might well be appealing to naturalistically inclined presentists, especially considering that presentists seem able to permit themselves only a naturalism with rather narrow limits. (But the narrowness of presentists' naturalism shows in dealing with Dowe's theory as well, as we shall see in the next section.) Third, the process view of causation appears better suited than discrete events-based views—such as the regularity, the counterfactual, and the primitivistic views—to host *causal realism*, i.e., the conception that causation is ultimately grounded on causal properties understood as bestowing dispositions to behave in certain ways upon the objects exemplifying them (see Chakravartty 2005: §4). Naturally, in the case of the CQ theory, it would be conserved quantities to play the role of the causally relevant properties at issue.[12] By endorsing causal realism, presentists can perhaps compensate for their antirealism towards the various cross-temporal relations allegedly involved in causation and thus meet the problem of causation in a way that is more satisfying to those eternalists who find their position better because it allows for a realistic stance towards those relations.[13]

[10] I thank the anonymous referee who has recommended that this issue be addressed and has also given some valuable suggestions to address it.

[11] See, e.g., part two of *The Oxford Handbook of Causation*.

[12] Note, however, that the claim that the process view is in accordance with causal realism is not defended by Dowe (who, in fact, seems to take no stance at all on the issue of causal realism).

[13] Naturalism and suitedness to host causal realism as merits of the CQ theory have been suggested by the anonymous referee mentioned in footnote 10.

3. The Spacetime Framework

Naturalism is an attractive feature of the CQ theory globally considered, and presentists may well agree on that. However, not every naturalistic element of the CQ theory is of benefit to presentism; in fact, two of them seem rather a threat to it.

As previously mentioned, the CQ theory is devised to allow for backwards causation (it is therefore time-symmetric and needs to be integrated by an explanation of the prevailing direction of causation). But it is very plausible that backwards causation is incompatible with presentism (see Faye 2015: §2); hence, it is very plausible that if backwards causation is real, then presentism is false. The reality of backwards causation, however, is rather controversial; hence, the inability to account for it does not appear to be a very disturbing inadequacy of presentism. Moreover, while this inability may represent a limit to presentists' naturalism (in the case that quantum mechanics really offers reasons to admit of backwards causation), at least it relieves them of the need to account for the idea that causation has a *prevalent* direction: presentists may take causation to have just one direction and causes to be definitionally prior to their effects (how to account for temporal precedence without invoking the precedence relation will be shown in §5).

A more serious problem for presentism, however, emerges from another, very noticeable naturalistic element of Dowe's theory: its being formulated drawing upon Minkowski's four-dimensional spacetime conception, i.e., the geometrical formulation of the special theory of relativity (STR) (see, in particular, Dowe 2000: 90-92). No doubt, this is a theoretical framework *prima facie* hostile to presentism. And that for at least two reasons: first, it is typical of eternalists to treat time as analogous to a fourth spatial dimension and to conceive of the universe as a *four-dimensional* block universe; second, as is well known, one of the most troubling arguments against presentism (and in favour of eternalism) is grounded on the relativisation of simultaneity at a distance involved by STR—or, more precisely, by its standard interpretation, i.e., the Minkowskian or Einsteinian one.

I think, however, that the *four-dimensionality* of Minkowski spacetime *per se* does not represent a problem for presentism: Galilean spacetime is also four-dimensional, but it raises no problem for presentism. By 'spacetime', one may mean a spatiotemporal entity, i.e., the concrete reality itself as it is conceived of by eternalists (the block universe); however, by 'spacetime' physicists primarily mean a *four-dimensional manifold*, i.e., a mathematical—hence, *abstract*—entity. Understood in this latter way, a spacetime can be considered neutral with regard to the various theories conflicting in temporal ontology: such a manifold can be used to *represent* an eternalist block universe or a presentist slice universe, depending on the section of the manifold we consider corresponding to what (tenselessly) exists (see Wüthrich 2013).

The real problem for presentism stems from the peculiar geometrical features of Minkowski spacetime as compared to the Galilean one. Minkowski spacetime, unlike the Galilean one, does not allow to define an invariant notion of simultaneity at a distance, which is a necessary prerequisite for an objective cosmically extended present. Since, I think, the notion of an objective cosmically extended present is essential to presentism, I consider as simply incoherent any attempt to reconcile presentism with STR by giving up this notion. Presentism needs a spacetime structure supporting absolute simultaneity. It is far beyond the scope of this paper to discuss whether such a "reactionary need" may be satis-

fied and, if so, how (by boldly rejecting STR as false, by endorsing a non-standard, Neo-Lorentzian interpretation of STR, or otherwise?); I shall therefore sidestep this problem and turn to the next one.

4. Causal Processes and Durative Events

As mentioned in §2, causal interactions are formed by the spatiotemporal overlapping of causal processes; causal processes are plausibly regarded as entities of a broader category, namely durative events; and durative events, in turn, are plausibly regarded as mereological sums of shorter and shorter durative events and, ultimately, of (infinite) instantaneous events ordered in temporal sequence. Two intertwined problems for presentism appear here: one concerns durative events' being *mereological sums* of shorter events, i.e., their being entities having *temporal parts*; the other concerns durative events' being *temporal sequences* of shorter events. In this section, I address the former problem (postponing the treatment of the latter one to the next section).

A durative event does not fit into the instantaneous present of the presentistic universe: if an event is a durative one, then it can be *only partly* present, i.e., only in some instantaneous part of itself. But the following principle, which might be called *Principle of Mereological Sums*, seems very plausible:

(MS) Necessarily, a mereological sum of parts $x_1 \ldots x_n$ (tenselessly) exists only if $x_1 \ldots x_n$ (tenselessly) exist.

(MS) can be justified by the following reason: the existence of a mereological sum of certain parts *conceptually* requires that *all* of its parts (tenselessly) exist; without any of its parts, something could not even qualify as a mereological sum *of them*. (Notice that I am not saying that by removing some part from a mereological sum, it would cease to be *a* mereological sum; I am saying that it would cease to be *that specific* mereological sum that is formed, among other parts, by *that* specific part.) So, if (MS) is true and there are durative events, then presentism is false.

I believe that, confronted with this problem, presentists must bite the bullet and give up the idea that there exist (tenselessly and thus present-tensedly) durative events. Admitting that is not too bad for presentists, though. Presentism can still allow for instantaneous events individually considered and can also allow for *temporal sequences* of them (temporal sequences that, of course, must be conceived of in some presentistically appropriate way: for example, in the way that will be expounded in the next section). And, most importantly, a (presentistically conceived of) sequence of instantaneous events would still be able to do, within the CQ theory, the basic theoretical work that is done by a causal process, namely "conveying" a conserved quantity from the cause to the effect.

While assuming this anti-realistic stance on durative events, and thus on causal processes and causal interactions too, presentists may decide to retain, for communicative ease only, *talk* of durative events, causal processes, and causal interaction. This may be done by setting some conditions for the usage of those locutions. We may convene (i) that if one or more objects exemplify through a period of time one or more attributes, then we *say* that a "durative event" occurs at that period of time; (ii) that if a physical object possesses a conserved quantity through a period of time, then we *say* that a "causal process" occurs at that period of time; (iii) that if two objects are spatially adjacent or overlapping through a

period of time during which they exchange a certain quantity, then we *say* that a "causal interaction" occurs at that period of time.[14] This metaphysically non-serious use of these locutions may be signalled by enclosing them between double quotation marks, as I have done just now. (How the talk of *periods of time* might be understood within a presentistic framework will be expounded in the next section.)

5. Cross-Temporal Relations

First, let us explain in a little more detail what cross-temporal relations are and why they are troublesome for presentism. As said in §1, a cross-temporal relation is, approximately, a relation that is exemplified by its relata at different times. More precisely, as Torrengo (2008: 15) writes:

> A relation R is cross-temporally exemplified by $x_1 \ldots x_n$ if and only if each x_i enters [i.e., exemplifies] R at a different time than some x_j.

So, if a cross-temporal relation is exemplified by one of its relata at the present time, then it is exemplified by at least one other relatum at a past time or at a future time. But, according to the so-called *Principle of Relations*,

(PR) necessarily, if $x_1 \ldots x_n$ exemplify a relation R, then $x_1 \ldots x_n$ (tenselessly) exist (each one at the time it exemplifies R).

Hence, any relatum that is cross-temporally related to a present entity *(tenselessly) exists at some non-present time*, and this is incompatible with presentism.

Dowe's CQ theory of causation, as can be seen from the concise exposition of it given in §2, appears to involve a variety of relations. In what follows, I shall review the ones that are cross-temporal or might be suspected to be so, and show how to dispose of those that really are. In line with the authors mentioned in §1, this disposal will be carried out by implementing a rephrasing strategy consisting in the replacement of any portion of causal discourse including predicates expressing the exemplification of a cross-temporal relation (e.g., 'causes' or 'is earlier than') by locutions able to play similar theoretical roles but having ontic commitments compatible with presentism, such as *primitive tensed properties*, i.e., properties incorporating tense (e.g. *having been true* and the like), and *simultaneously exemplified relations*, or *simultaneous relations* for short, i.e., roughly, relations exemplified by all of their relata at the same time.[15] Thus, an alternative way of expressing the CQ theory will be shown, one that does not force us admitting of cross-temporal relations.

5.1 Precedence Relation and Handling of Tense

The *precedence* (or *earlier than*) relation does not play a showy role in the CQ theory. It is involved, nevertheless: the CQ theory presupposes that, typically (alt-

[14] As regards the counterintuitive idea that an object may not simply be spatially adjacent with another but may also spatially overlap with it, we must be reminded that, in Dowe's CQ theory, by 'objects' are meant not only the material objects of common sense, such as chairs and animals, but also the impalpable objects of physics, such as fields and waves, for which it makes sense talking about spatial co-location and, thus, of spatial overlapping.

[15] More precisely, a relation R is simultaneously exemplified by $x_1 \ldots x_n$ if and only if each x_i exemplifies R at the same time as any other.

hough not necessarily), causes are earlier than their effects; moreover, as pointed out in §4, presentists' "causal processes", while not being (real) causal processes, still consist in *temporal sequences* of instantaneous events. So, presentists must account both for the temporal precedence of causes over their effects and for "causal processes" in some way that does without the precedence relation. This can be achieved by resorting to metric tense operators metaphysically interpreted in a certain way.

Let us first introduce *non-metric tense operators*: the past tense operator '**P**', reading 'it was true that', and the future tense operator '**F**', reading 'it will be true that'.[16] These operators are the basic expressive devices to which, typically, presentists resort in order to attain a presentistically suitable interpretation of sentences containing past-tensed or future-tensed predication. Such an interpretation is carried out in two steps: first, the past- or future-tensed sentence at issue is reformulated in the present tense; then, the apt tense operator is placed in front of the thus obtained sentence. Consider, for example, a past-tensed sentence and a future-tensed sentence expressing some generic instantaneous event of the sort involved in Dowe's theory of causation:

(1) Something had a certain amount of quantity q,

and

(2) Something will have a certain amount of quantity q.

Assuming a tense operator-based semantics, (1) and (2) will be rendered, respectively, as:

(1.1) **P**(something has a certain amount of quantity q),

and

(2.1) **F**(something has a certain amount of quantity q),

whose structures may be further made explicit, respectively, as:

(1.2) $\mathbf{P}(\exists y \exists z\,(q(y) = z))$,

and

(2.2) $\mathbf{F}(\exists y \exists z\,(q(y) = z))$.

Here, Dowe's symbolical rendering for the having (or possession) of a quantity is adopted (the only difference being that the variable ranging over possible amounts of the quantity at issue—here, the variable z, ranging over possible amounts of q—is bound by the quantifier, whereas in Dowe's original formulation it is left free). For a better readability, I shall use different individual variables, possibly with subscripts, for different kinds of entities: y, y_1, etc. for objects; z, z_1, etc. for amounts of quantities; later on, I shall also use x, x_1, etc. for another sort of entity: intervals of time or, better, degrees of tensedness.

The idea that, from a metaphysical point of view, underlies the presentistic appeal to a tense operator-based semantics for past-tensed and future-tensed predication is, it seems to me, that it allows to qualify as past or future an event (e.g., the possession of an amount of quantity q by some physical object) in an

[16] More precisely, '**P**' and '**F**' are called *weak* (non-metric) tense operators, in contrast with *strong* (non-metric) tense operators, namely '**H**' and '**G**', which read, respectively, 'it has always been true that' and 'it will always be true that'. On the semantics of tense, see Ludlow (2006); on tense logic, see Galton and Goranko (2015).

"indirect" way, namely by saying (in the metalanguage) that the present-tensed proposition representing it (in our example, the one expressed by the present-tensed claim 'something has a certain amount of quantity *q*') was or will be true. This supposedly enables presentists to claim that something *was* or *will be* in certain way, *existed* or *will exist*, without entailing that that thing *is tenselessly* in that way *at some past or future time* or *tenselessly exists at some past or future time*–which is something presentists deny (and, vice-versa, eternalists affirm).[17] So, presentists are supposedly able to claim that (1) and (2), interpreted as (1.1)/(1.2) and (2.1)/(2.2), may be true without entailing the truth of 'something (tenselessly) has a certain amount of quantity *q* at some past or future time' and the truth of 'something (tenselessly) exists at some past or future time'. Of course, tense operators must be primitive to do the semantic work that they are supposed to do within a presentistic framework.[18]

But, I think, the nature of tense operators needs to be specified a little further. The problem of a presentistically suitable interpretation of past-tensed and future-tensed predication re-emerges for the metalinguistic renditions of '**P**' and '**F**' ('it was true that' and 'it will be true that'), which do in turn contain past-tensed and future-tensed predications: surely, by claiming (in the metalanguage) that a certain present-tensed proposition was or will be true, presentists do *not* want to entail that that proposition *is (tenselessly) true at some past or future time*: this would in turn entail that that proposition *(tenselessly) exist at that time*, which, again, is something that presentists deny. I think that a viable way for presentists to fix this problem may be to further reduce past-tensed and future-tensed predications of the property *being true* to present-tensed predications of *primitive tensed properties*, i.e., respectively, of the past-tensed property *having been true* and of the future-tensed property *going to be true* (as is known, properties of this kind, somehow "incorporating" tense in themselves, are often invoked by presentists in dealing with the grounding problem: see, e.g. Bigelow 1996).[19] If this suggestion is accepted, then, (1.2) and (2.2) will be (metalinguistically) understood as follows:

(1.3) The proposition $[\exists y \exists z\,(q(y) = z)]$ has *having been true*,

and

(2.3) The proposition $[\exists y \exists z\,(q(y) = z)]$ has *going to be true*.

It is now possible to introduce metric tense operators. A *metric tense operator* is a tense operator that not only tells that a proposition was or will be true but also

[17] E.g., a sentence like (1) may be rendered in eternalistic terms as 'There (tenselessly) exists something that (tenselessly) has a certain amount of quantity at a past time', where the adjective 'past' may be understood in turn as 'that is (tenselessly) earlier than the present time'.

[18] An anonymous referee has complained that the employment of a tenseless predication of existence in defining presentism is in tension with the ideological commitment to primitive tense operators. I must disagree with her/him on this point. Tenseless predication and recourse to primitive tense operators are compatible and are both of use to presentists: tenseless predication is required to capture the sense in which the present, unlike the past and the future, *exists*; primitive tense operators are required to capture the sense in which the past and the future, respectively, *existed* and *will exist*.

[19] The idea that tense operators may be understood, ultimately, as present-tensed predications of primitive past- or future-tensed properties has emerged in conversation with Francesco Orilia.

specifies *when* it was or will be true by indicating the amount of time intercurrent between the present instant and the instant at which the proposition was or will be true. The metric past tense operator is '$\mathbf{P}_x$', which reads 'it was true the interval x ago that…'; the metric future tense operator is '$\mathbf{F}_x$', reading 'it will be true the interval x hence that…'. The variable x ranges over positive real numbers, which, chosen a suitable unit of measure (e.g., seconds, hours, days), are taken to specify the sizes of time periods between past or future instants and the present instant.

As they stand, however, metric tense operators are not immediately acceptable to presentists: tense operators resort to *periods of time*, which are something that presentists cannot accept in their ontic inventory, and that for exactly the same reasons for which they cannot accept durative events: presumably, periods of time are composed of shorter and shorter periods of time, and ultimately of infinite durationless instants, ordered by the precedence relation. So, what can we do? We could conceive of the metric indices as specifying not time periods but *degrees of past-tensedness* (of the property *having been true*) and *degrees future-tensedness* (of the property *going to being true*), taken, again, as primitive. We could think that the past-tensedness (of *having been true*) expressed by '**P**' and the future-tensedness (of *going to be true*) expressed by '**F**' have degrees, and that these degrees are picked out by the corresponding metric tense operators. We can interpret this in the terms of the distinction between determinable properties and determinate properties: while '**P**' expresses simply the exemplification of the determinable property of *having been true*, '$\mathbf{P}_x$' expresses the exemplification of a determinate *having been true*, i.e., one having a determinate amount of past-tensedness, namely x; and similar considerations hold, *mutatis mutandis*, for the future tense operators. Again, presentists may retain *talk* of "periods of time" for its usefulness, with the understanding, however, that what metaphysically underlies this talk, from their perspective, are determinate past-tensed and future-tensed properties of propositions, and not real, temporally extended periods of time.

In order to simplify our symbolic rendering, we may introduce a single tense operator '$\mathbf{T}_x$' (for 'Tensedness') and, assuming that P stands for any present-tensed sentence and x ranges over real numbers, convene that: for $x < 0$, '$\mathbf{T}_xP$' is equivalent to '$\mathbf{P}_{|x|}P$'; for $x = 0$, '$\mathbf{T}_xP$' is equivalent to P; and for $x > 0$, '$\mathbf{T}_xP$' is equivalent to '$\mathbf{F}_xP$' (see Galton and Goranko 2015: §7.2). So, by setting a negative value for x, we (present-tensedly) predicate, of the proposition $[P]$, the property *having been true* with a degree x of past-tensedness; by setting a positive value for x, we predicate, of $[P]$, the property *going to be true* with a degree x of future-tensedness; and by setting value 0 for x, we do not predicate, of $[P]$, any tensed property.

Equipped with metric tense operators understood in this way, we can express the temporal precedence of an event over another without employ the predicate 'is earlier than' (or its synonyms). E.g., instead of writing 'the event $\exists y_1\exists z_1(q(y_1) = z_1)$ is earlier than the event $\exists y_2\exists z_2(q'(y_2) = z_2)$',[20] we can write as follows:

(3) $\exists x_1\exists x_2((x_1 < x_2) \wedge (\mathbf{T}_{x1}(\exists y_1\exists z_1(q(y_1) = z_1)) \wedge \mathbf{T}_{x2}(\exists y_2\exists z_2(q'(y_2) = z_2))))$.

The idea behind this paraphrase (and the others that will follow) is that any part of discourse expressing the exemplification of the precedence relation can

[20] Or, mindful of what has been said in footnote 9, '$[\![\exists y_1\exists z_1(q(y_1) = z_1)]\!]$ is earlier than $[\![\exists y_2\exists z_2(q'(y_2) = z_2)]\!]$'.

be replaced by a *conjunction of tense-logical propositions*: propositions are taken to exist *at the present time*; consequently, the relation expressed by '∧', and by any other dyadic propositional connective, may be taken to relate two present entities and, then, to be simultaneously exemplified;[21] each proposition, however, including a metric tense operator, is in turn able to "indirectly" qualify a certain event as past or future (in a certain degree) or as present.

By means of metric tense operators, we can also describe a temporal sequence of instantaneous events constituting a "causal process" in a way that is fully compatible with presentism. E.g., a "causal process" consisting in the possession of the quantity q by a generic object y, beginning m units of time ago or hence and ending n units of time ago or hence, may be expressed as follows:

(4) $\forall x\,((m \leq x \leq n) \rightarrow \mathbf{T}_x(\exists y \exists z\,(q(y) = z))$,

which is equivalent to a very long—in fact, infinite—disjunction of all sentences of the form '$\mathbf{T}_i(\exists y \exists z\,(q(y) = z)$' with $m \leq i \leq n$, i.e., something like the following:

(4.1) $\mathbf{T}_m(\exists y \exists z\,((q(y) = z) \wedge \ldots \wedge \mathbf{T}_n(q(y) = z)))$.

It should be noticed that (4) as well as (4.1) allow z to take different values through time. In fact, according to Dowe, for a process to be causal, it is only necessary that the constitutive object of the process possesses a conserved quantity at every stage of it, no matter whether the amount of the quantity remains constant throughout the process or not. And the same, of course, must hold for "causal processes" as well. To designate concisely such "causal processes" in which the amount of the quantity is not stable through time, we may use Dowe's original symbology: e.g., in the case of the "causal process" represented by (4) and (4.1), we may write '$\Delta q(y)$'.

Obviously, the tense operator-based account of temporal sequences adopted in (4) and (4.1) can be extended, with the apt changes, to "durative events" that are not "causal processes" (e.g., a leaf's being green for two hours).[22] And it can be extended, again with the apt changes, to "causal interactions" as well, as will be shown in a moment.

5.2 Exchange Relation

A quantity exchange may be taken to consist in the exemplification of a triadic *exchange* relation having a quantity and two objects as relata; correspondingly, a sentence like 'two objects (presently) exchange the quantity q' may be symbolised by a formula such as '$\exists y_1 \exists y_2 (E_q y_1 y_2)$'.[23] Of course, other metaphysical accounts and symbolisations of the structure of quantity exchanges are feasible. The important issue, here, is whether the exchange relation is cross-temporal or simultaneous. In this connection, I think, two relevant cases should be distinguished: the case where two objects exchange some quantity *in a direct way*, i.e., without exchanging it with

[21] I think that, from a presentistic perspective, propositions must be understood as temporal (perhaps omnitemporal) entities, not as atemporal ones; and that because they must be able to change truth-value across time (as time flows), whereas atemporal entities are subtracted to the possibility of any sort of change.

[22] For a different way to account for "durative events" within presentism see Orilia (2012).

[23] Thus, an exchange involving more than two objects or more than one quantity must be understood as involving *different exemplifications* of the exchange relation (not a single exemplification with more than two objects or more than one quantity).

some intermediary object (e.g., the exchange of linear momentum between two billiard balls colliding with each other), and the case where they exchange the quantity *in an indirect way*, i.e., by exchanging it with some intermediary object (e.g., the exchange of energy between the Sun and the Earth by means of photons traveling from the former to the latter). If two objects exchange a quantity without intermediary objects, the exchange occurs in both objects at the same time: the giving up of a quantity by one object temporally coincides with the gaining of it by the other object; there is no temporal delay between the giving up and the gaining of the quantity.[24] And this suggests that, where the exchange is direct, the exchange relation is exemplified simultaneously. It is important to point out a necessary condition for an exchange to be direct: an exchange is direct only if it is *local*, i.e., if, throughout the exchange, the objects involved in it are spatially adjacent or overlapping. On the other hand, if the exchange is indirect, it is possible that the periods during which each of the two objects exchanges the quantity with the other do not coincide: surely, this happens in the case in which the exchange is *non-local*, i.e., it is between two spatially non-adjacent or non-overlapping objects: while locality is a necessary condition for an exchange to be direct, non-locality is a sufficient condition for an exchange to be indirect. In such a case, the intermediary objects behave like "vehicles" of the quantity: they move, thereby conveying the quantity (like in the previously mentioned example of the Sun exchanging energy with the Earth). But moving takes time; hence, an indirect non-local exchange between two objects involves some temporal delay between the giving up of the quantity by one object and the gaining of it by the other object. It would seem, thus, that an indirect and non-local exchange of a quantity involves a cross-temporal exemplification of the exchange relation.

Quite clearly, the exchanges *primarily* involved in causal interactions (as they are understood within the CQ theory) are those of *direct* (and, thus, local) kind. A causal interaction between two causal processes, therefore, may be viewed as a *simultaneous exemplification of the exchange relation* (by the quantity and the objects constituting the interacting causal processes) *lasting for a certain time*, which may in turn be understood as a *temporal sequence* of instantaneous simultaneous exemplification of the exchange relation (by the quantity and the objects constituting the interacting causal processes).[25] Presentists may understand a "causal interaction" between two "causal processes" in an analogous manner, provided that the notion of temporal sequence is understood, again, in a suitable way—for example, by means of metric tense operators. The simplest case of a "causal interaction", say, the case where two generic objects, y_1 and y_2, exchange the quantity q—something that Dowe would represent by a formula like '$\Delta q(y_1)$, $\Delta q(y_2)$'—could be rendered in a manner that fits presentism as follows:

(5) $\forall x\,((m \leq x \leq n) \rightarrow \mathbf{T}_x(\exists y_1 \exists y_2(E_q y_1 y_2)))$,

where m and n are the intervals of time intercurrent between the present instant and, respectively, the beginning and the end of the "causal interaction" or, better, the corresponding degrees of tensedness.

[24] This is what I have been able to glean from physics textbooks with the help of friends who have an understanding of physics much deeper than mine.

[25] Of course, this neither amounts to say nor entails that an exchange relation might be exemplified for just one instant: it is exemplified so long as the two objects undergo a change in amount of the relevant quantity.

However, it should be noted that it may well be—and, in fact, it is commonly—the case that direct exchanges between two objects, in turn, involve indirect non-local exchanges between *parts of those objects*. During a direct exchange of a quantity between two objects, their parts also exchange that quantity; but any two spatially non-adjacent or non-overlapping parts can exchange it only in an indirect way. E.g., if two billiard balls collide, each ball, considered in its entirety, exchanges with the other, considered in its entirety, a certain amount of momentum in a direct way (through the adjacent parts); however, during the collision, the various parts of each ball are also involved in exchanges of momentum and, of course, any two spatially non-adjacent or non-overlapping parts can exchange it only in an indirect way. As previously mentioned, indirect exchanges involve a temporal delay and seem thus to involve a cross-temporal exemplification of the exchange relation. To avoid this, where the predicate 'exchanges' is intended to express an indirect exchange, we may replace it with a presentistically apt description the *chain of direct exchanges* between the intermediary spatially adjacent or overlapping objects conveying the quantity, i.e., by a conjunction of formulae like (5). (Given the impractical verbosity of such conjunctions, presentists may nevertheless keep using the predicate 'exchanges' also to express indirect exchanges, provided that it is considered simply as a useful, but metaphysically inappropriate, way of talking.)

5.3 Causal Relation

Lastly, we have the cross-temporal relation that has given rise to the discussion about presentism and causation in the first place: the causal relation itself. The causal relation may be dispensed with by replacing any causal claim containing the troublesome predicate 'causes' (or its equally troublesome companions, such as 'is causally connected to' etc.), e.g., 'the event $\exists y_1 \exists z_1(q(y_1) = z_1)$ causes the event $\exists y_2 \exists z_2(q'(y_2) = z_2)$',[26] by a causal claim of the following form:

(6) $\exists y_2 \exists z_2(q'(y_2) = z_2)$ because $\exists y_1 \exists z_1(q(y_1) = z_1)$,

where an instance of causation is rendered by means of the dyadic connective '*because*', which expresses a simultaneous relation holding between propositions representing the effect-event and the cause-event.[27] Of course, formulae like (6) must be integrated by tense operators to express the temporal priority of the cause-event over the effect-event, for example, as follows:

(6.1) $\exists x_1 \exists x_2((x_1 < x_2) \wedge (\mathbf{T}_{x2}(\exists y_2 \exists z_2(q'(y_2) = z_2)) \text{ because } \mathbf{T}_{x1}(\exists y_1 \exists z_1(q(y_1) = z_1))))$.

6. Conclusion

In the light of what has been said in the previous sections, Dowe's analysis of the grounds of causation may be reformulated as follows. A generic instance of causation, e.g.:

(7) $\exists x_1 \exists x_2(\mathbf{T}_{x2}(\exists y_2 \exists z_2(q'(y_2) = z_2)) \text{ because } \mathbf{T}_{x1}(\exists y_1 \exists z_1(q(y_1) = z_1)))$

[26] Or, perhaps better, '$[\![\exists y_1 \exists z_1(q(y_1) = z_1)]\!]$ causes $[\![\exists y_2 \exists z_2(q'(y_2) = z_2)]\!]$'.

[27] An operator-based account of this sort has been originally suggested as a presentistically suitable account of the primitivistic view of causation. However, it can be assumed as a general replacement for the problematic causal claims of the typical form '*c* causes *e*' (where *c* and *e* stand for events).

holds if and only if the following four conditions are satisfied:

(i) $x_1 < x_2$;

(ii) $\forall x((x_1 \leq x \leq x_2)\ (\mathbf{T}_x(q(y_1) = z_1) \vee \mathbf{T}_x(q'(y_2) = z_2) \vee \mathbf{T}_x(\mathrm{E}_q y_1\, y_2) \vee \mathbf{T}_x(\mathrm{E}_{q'} y_1\, y_2)))$;

(iii) any change of object from y_1 to y_2 and any change of conserved quantity from q to q' in the temporal succession described by (ii) occur at a "causal interaction" involving the following changes: $\Delta q(y_1)$, $\Delta q(y_2)$, $\Delta q'(y_1)$, $\Delta q'(y_2)$; and

(iv) for any exchange in (iii) involving more than one conserved quantity, the changes in quantities are governed by a single law of nature.

Condition (i) expresses the temporal precedence of the cause-event over the effect-event in the form of a relation between degrees of tensedness (as previously mentioned, given up the possibility of backwards causation, presentists may assume temporal precedence as the basis for the distinction between cause and effect); condition (ii) expresses the temporal sequence of instantaneous possession-events and instantaneous exchange-events, i.e., the "causal processes" and the "causal interactions" going from the cause-event to the effect-event; conditions (iii) and (iv) are simply reformulations with minimal adjustments of the conditions (1) and (2) given in Dowe's analysis.

In this paper, I have been engaged in the attempt of reformulating the CQ theory in a way that is compatible with presentism. This has been done by replacing a series of interconnected, presentistically unwelcome ontic commitments—in first place, to cross-temporal relations and to non-present events, objects, and times, but also to (real) causal processes and (real) causal interactions—with presentistically acceptable ontic commitments, namely to propositions, simultaneous relations, and primitive tensed properties. The most controversial of these ontic commitments is doubtlessly the third one, which is at the same time an ideological commitment. In fact, the plausibility of primitive tensed properties has been challenged by Sider (2001: 36-41) and Cameron (2011); so, the plausibility of the presentistic account of the CQ theory put forward in this paper ultimately depends on whether their objections can be properly answered. Carrying out a defence of primitive tensed properties, however, is a task that must be left for another time.[28]

References

Armstrong, D. 1997, *A World of States of Affairs*, Cambridge: Cambridge University Press.

Bigelow, J. 1996, "Presentism and Properties", *Philosophical Perspectives*, 10, 35-52.

Bourne, C.A. 2006, *A Future for Presentism*, Oxford: Oxford University Press.

Brogaard, B. 2006, "Tensed Relations", *Analysis*, 66, 194-202.

[28] I wish to thank Francesco Orilia, Nathan L. Oaklander, Michele Paolini Paoletti, and the (many) anonymous referees for their valuable suggestions on previous versions of this paper; I also wish to thank Erica Chiaverini, Antonio De Grandis, and Davide De Grandis for the very useful conversations about the topics in physics addressed in the paper (of course, I am solely responsible for any misunderstanding or mistake possibly contained in it).

Brogaard, B. 2013, "Presentism, Primitivism and Cross-Temporal Relations: Lessons from Holistic Ersatzism and Dynamic Semantics", in Ciuni, Torrengo and Miller 2013, 253-77.

Cameron, R. 2011, "Truthmaking for Presentists", in Bennett, K. and Zimmerman, D.W. (eds.), *Oxford Studies in Metaphysics*, Vol. 6, Oxford: Oxford University Press, 55-101.

Casati, R. and Varzi, A. 2006, "Events", in Zalta E.N. (ed.), *The Stanford Encyclopedia of Philosophy*, Summer 2014 edition.

Castañeda, H.-N. 1984, "Causes, Causity, and Energy", in French, P., Uehling, T. and Wettstein, H. (eds.), *Midwest Studies in Philosophy*, IX, Minneapolis: University of Minnesota Press, 17-27.

Chakravartty, A. 2005 "Causal Realism: Events and Processes", *Erkenntnis*, 63, 7-31.

Ciuni, R., Torrengo, G. and Miller, K. (eds.) 2013, *New Papers on the Present: Focus on Presentism*, München: Philosophia Verlag.

Crisp, T. 2005, "Presentism and 'Cross-temporal' Relations", *American Philosophical Quarterly*, 42, 5-17.

Dowe, P. 2000, *Physical Causation*, Cambridge: Cambridge University Press.

Fair, D. 1979 "Causation and the Flow of Energy", *Erkenntnis*, 14, 219-50.

Faye, J. 2015, "Backwards Causation", in Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy*, Autumn 2015 edition.

Galton, A. and Goranko, V. 2015, "Temporal Logic", in Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2015 edition.

Hestevold, H.S. and Carter, W.R. 2002, "On Presentism, endurance, and Change", *Canadian Journal of Philosophy*, 32, 491-510.

Kistler, M. 2006, *Causation and Laws of Nature*, London: Routledge.

Ludlow, P. 2006, "Tense", in Lepore, E. and Smith, B. (eds.), *The Oxford Handbook of Philosophy of Language*, Oxford: Oxford University Press, 689-715.

McDaniel, B. 2010, "Presentism and Absence Causation: An Exercise in Mimicry", *Australasian Journal of Philosophy*, 88, 323-32.

Orilia, F. 2012, "Dynamic Events and Presentism", *Philosophical Studies*, 160, 407-14.

Salmon, W.C. 1997, "Causality and Explanation: A Reply to Two Critiques", *Philosophy of Science*, 64, 461-77.

Salmon, W.C. 1998, *Causality and Explanation*, Oxford: Oxford University Press.

Sider, T. 1999, "Presentism and Ontological Commitment", *Journal of Philosophy*, 96, 325-47.

Sider, T. 2001, *Four-Dimensionalism: An Ontology of Persistence and Time*, New York: Oxford University Press.

Torrengo, G. 2008, *Time and Cross-temporal Relations*, Milano: Mimesis.

Torrengo, G. 2012, "Time and Simple Existence", *Metaphysica*, 13, 125-30.

Wüthrich, C. 2013, "The Fate of Presentism in Modern Physics", in Ciuni, Torrengo and Miller 2013, 91-131.

Indicative Conditionals as Strict Conditionals

Andrea Iacona

University of Turin

Abstract

This paper is intended to show that, at least in a considerably wide class of cases, indicative conditionals are adequately formalized as strict conditionals. The first part of the paper outlines three arguments that support the strict conditional view, that is, three reasons for thinking that an indicative conditional is true just in case it is impossible that its antecedent is true and its consequent is false. The second part of the paper develops the strict conditional view and defends it from some foreseeable objections.

Keywords: Indicative conditional, Strict conditional, Material conditional, Negation.

1. Preliminary Clarifications

Let us assume that > stands for 'if then' as used in indicative sentences, and that 1 and 0 designate truth and falsity. According to the strict conditional view, the truth conditions of $p > q$ are defined relative to a possible world w as follows:

Definition 1: $[p > q] = 1$ in w if and only if, for every w', either $[p] = 0$ in w' or $[q] = 1$ in w'.

As is natural to expect, the set of possible worlds over which 'every' ranges may vary from context to context, just as in any other quantified sentence. To say that $p > q$ is true simpliciter is to say that $p > q$ is true in the actual world.

As the initial assumption about > implies, this paper focuses on indicative conditionals. From now on we will take for granted that 'conditional' abbreviates 'indicative conditional'. This is not to suggest that the case of counterfactuals is essentially different. On the contrary, most of what will be said about conditionals can be extended, *mutatis mutandis*, to counterfactuals. But for the sake of simplicity we will not deal with such extension.[1]

Moreover, the symbol > is not intended to characterize a semantically homogeneous class of sentences. Conditionals may be used in more than one way,

[1] Iacona 2015 deals with counterfactuals.

Argumenta 4, 1 (2018): 177-192
ISSN 2465-2334

© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.iac

and it is reasonable to expect that different criteria of assessment are appropriate in different cases. In particular, we will take for granted that, although in some cases it is plausible to read $p > q$ as $p \supset q$, that is, as a material conditional, in other cases it is not. From now on, the label 'nonmaterial' will be used generically for any conditional that at least prima facie is not tractable as a material conditional. The literature on conditionals mostly focuses on nonmaterial conditionals, and we will do the same.

The next three sections outline three arguments that support the strict conditional view. The three arguments hinge on three observations that will be called 'facts'. This is not to say that they are truths written in stone. Perhaps there are no such truths. Or at least, the history of the debate on conditionals shows that everything—or almost everything—can be questioned. So, none of the three arguments is intended to provide a conclusive reason to accept definition 1. Still, each of them deserves attention.

2. First Argument

The first argument hinges on the following observation:

Fact 1: It seems that in some cases one can assert $p \wedge q$ but deny $p > q$.

Although in many cases $p \wedge q$ and $p > q$ are both assertable, it seems that nothing in principle prevents us from thinking that one can accept $p \wedge q$ but reject $p > q$. On the assumption that the rejection of $p > q$ justifies the acceptance of $\sim (p > q)$, this is to say that there are cases in which $p \wedge q$ and $\sim (p > q)$ are both assertable. Suppose that a coin is to be tossed twice and I bet that it will come up heads both times. Consider the following sentence, uttered just after the bet and before the first toss:

(1) If at least one head will come up, I will win.

In this case it seems correct to deny (1), and the appearance of correctness of this denial does not vanish if the coin comes up heads both times. Yet in that case it would be correct to say that the antecedent and the consequent of (1) were both true.[2]

Fact 1 raises a controversial issue. Some theorists of conditionals think that $\sim (p > q)$ is inconsistent with $p \wedge q$ because they believe in *conjunction conditionalization*, that is, they believe what follows:

(CC) $p \wedge q$ entails $p > q$.

For example, Edgington says:

> Establishing that the antecedent and consequent are true is surely one incontrovertible way of verifying a conditional.[3]

Clearly, if $p \wedge q$ entails $p > q$, then $p \wedge q$ and $\sim (p > q)$ form an inconsistent set. However, CC cannot be invoked to dismiss fact 1 unless it is independently justified. Presumably, if one takes fact 1 at face value, it is because one thinks that the assertability of $p > q$ requires that a certain relation obtains between p and q.

[2] The example is adapted from McDermott 2007.
[3] Edgington 1986: 24.

But if one thinks so, then one may coherently deny CC, because p and q may be assertable even if that relation does not obtain.

One way to see that CC is dubious is to think that CC entails *conjunction bi-conditionalization*, that is:

(CB) $p \wedge q$ entails $(p > q) \wedge (q > p)$.

CB can be derived from CC by using the standard rules of *conjunction introduction* (CI) and *conjunction elimination* (CE):

1 [1] $p \wedge q$	A
1 [2] $p > q$	CC1
1 [3] q	CE1
1 [4] p	CE1
1 [5] $q \wedge p$	CI 3,4
1 [6] $q > p$	CC 5
1 [7] $(p > q) \wedge (q > p)$	CI 2,6

But it is easy to see that CB has implausible instances. For example, the following sentence seems to differ from (1) because it is clearly true:

(2) If I will win, then at least one head will come up.

However, according to CB there is no difference between (1) and (2), because (1) and (2) are both true, assuming that their constituents are true. Here is another example. I always use my bicycle to move around, and I cycle to work whenever I can. When it is a beautiful sunny day, of course, riding my bicycle is especially enjoyable. But also when it is cold, cloudy, or raining, I still prefer cycling than walking. I refrain from cycling only in case of ice or snow storm. Now suppose that I am in my office, and that two colleagues of mine, who have not looked out of the window since their arrival, utter the following conditionals:

(3) If it is a beautiful sunny day, then he came by bicycle
(4) If he came by bicycle, then it is a beautiful sunny day

In this situation it seems that one of them is right and the other is wrong. However, if it is actually a beautiful sunny day and I came by bicycle, CB entails that (3) and (4) are both true.[4]

The first argument rests on the assumption that fact 1 is a datum that deserves an explanation, so it is an argument for those who do not have a strong faith in CC. On this assumption, it turns out that the strict conditional view is more credible than other theories of conditionals, because it provides a better account of fact 1.

The strict conditional view explains fact 1 as follows: in some cases it seems that one can assert $p \wedge q$ but deny $p > q$ because in those cases the former is true but the latter is false. If $[p] = 1$ in w and $[q] = 1$ in w but there is a w' such that $[p] = 1$ in w' and $[q] = 0$ in w', $[p \wedge q] = 1$ in w but $[p > q] = 0$ in w. For example, in the case of the coin one can say that (1) is false, because there are possible worlds in which at least one head comes up and I do not win.

Now we will examine four well known alternatives to the strict conditional view, in order to show that none of them can explain fact 1. The first is the truth-functional view, the view that we find in every logic textbook:

Definition 2: $[p > q] = 1$ if and only if either $[p] = 0$ or $[q] = 1$.

[4] Butcher 1983: 89-90, takes the fact that CC entails CB as a reason against CC.

As long as assertability conditions are understood as truth conditions, the truth-functional view is patently inadequate to explain fact 1. Since definition 2 entails that $[p > q] = 0$ if and only if $[p] = 1$ and $[q] = 0$, it turns out that one cannot assert $p \wedge q$ but deny $p > q$.

Of course, it is not essential to the truth-functional view that assertability conditions are understood as truth conditions. If one draws a distinction between truth and assertability, one can maintain definition 2 and claim that fact 1 is to be explained in terms of assertability rather than in terms of truth. But then the question to be addressed concerns the account of the assertability conditions of $p > q$ rather than the truth-functional view itself.[5]

The second alternative is the probabilistic view, which is based on Adams's idea that the assertability of $p > q$ can be described in terms of conditional probability. If $P(p)$ is the probability of p and $P(q|p)$ is the probability of q given p, then $P(p > q) = P(q|p)$. So, $p > q$ has an epistemic value defined as follows:

Definition 3: $[p > q] = P(q|p)$.

Here it is assumed that $P(q|p) = P(p \wedge q)/P(p)$ for $P(p) > 0$. The values 1 and 0 correspond to the maximum and the minimum degree of belief. Although Adams does not provide a semantics for compound sentences with conditionals as parts, it is consistent with his idea to assume that definition 3 can be combined with the following principle:

(N) $[\sim (p > q)] = 1 - [p > q]$.

Definition 3 and (N) entail that $[\sim (p > q)] = 1 - P(q|p)$. Since $1 - P(q|p) = P(\sim q|p)$, given that $P(q|p) + P(\sim q|p) = 1$, it follows that $[\sim (p > q)] = [p > \sim q]$.[6]

The probabilistic view can be phrased in at least two ways. The simplest way—perhaps the closest to Adams's proposal—is to say that a sentence is assertable when its value is greater than the value of its negation, that is, when its probability is greater than 0.5. This version of the view squares ill with fact 1, as it entails that $p \wedge q$ and $\sim (p > q)$ cannot both be assertable. Suppose that (a) $[\sim (p > q)] > 0.5$ and (b) $[p \wedge q] > 0.5$. From (a) and (N) we get that $1 - [p > q] > 0.5$. By definition 3, this means that $1 - P(q|p) > 0.5$. Since $P(\sim q|p) = 1 - P(q|p)$, we get that $P(\sim q|p) > 0.5$. It follows that $P(q|p) \leq 0.5$. But $P(q|p) = P(p \wedge q)/P(p)$, so $P(p) = P(p \wedge q)/P(q|p)$. From this and (b) we get that $P(p) > 1$, which is absurd.

The second way to phrase the probabilistic view is slightly more sophisticated, in that it leaves room for contextual variation: $p > q$ is assertable in a context if and only if the conditional probability of q given p is greater than a number n fixed by the context, where $n > 0.5$. This second version of the view makes no significant advance as far as fact 1 is concerned. An inconsistency result can still be obtained by generalizing the reductio outlined above. Suppose that (a) $[\sim (p > q)] > n$ and (b) $[p \wedge q] > n$. From (a) and (N) we get that $1 - [p > q] > n$. By definition 3, this means that $1 - P(q|p) > n$. Since $P(\sim q|p) = 1 - P(q|p)$, we

[5] Lewis 1976, Jackson 1979, and others endorse definition 2 but claim that the assertability of $p > q$ can be measured in terms of probability. So the discussion of the second alternative applies also to such proposals.

[6] Adams 1965 outlines the view. Adams explicitly claims that $\sim (p > q)$ is equivalent to $p > \sim q$. In particular, Adams 1968: 271 presents a metalinguistic definition of $\sim (p > q)$ as an abbreviation of $p > \sim q$. Stalnaker 1970 also outlines a probabilistic semantics which combines definition 3 with (N).

get that $P(\sim q|p) > n$. It follows that $P(q|p) \leq n$. But $P(q|p) = P(p \wedge q)/P(p)$, so $P(p) = P(p \wedge q)/P(q|p)$. From this and (b) we get that $P(p) > 1$, which is absurd.[7]

The third alternative is the belief revision view, which has been elaborated by Gärdenfors and others. On this view, conditionals are defined as acceptable relative to belief states, understood as deductively closed sets of sentences. Let f be a belief revision function, that is, a function that, for a belief state K and a sentence p, gives us a revised belief state $f(K, p)$. The acceptability conditions of $p > q$ are given as follows:

Definition 4: $[p > q] = 1$ relative to K if and only if $q \in f(K, p)$.

Here 1 indicates acceptance, while 0 indicates non-acceptance or rejection: $[p > q] = 0$ relative to K if and only if it is not the case that $q \in (K, p)$. Since (N) holds, we get that $[\sim (p > q)] = 1$ relative to K if and only if $[p > q] = 0$ relative to K. To say that $[p > q] = 1$ relative to K is to say that there is a deductively closed set of sentences $s(K)$ which includes K and $p > q \in s(K)$. $s(K)$ is understood as a support set, that is, as a set whose acceptability is grounded on the adoption of K. The distinction between K and $s(K)$ matters only for conditionals, as it is assumed that membership in K means full belief while the acceptance of a conditional does not amount to full belief. For any p that does not contain >, if $p \in s(K)$ then $p \in K$. In this framework, a conditional can be defined as assertable for a speaker when it is acceptable relative to the speaker's belief state.[8]

The belief revision view does not explain fact 1. Given any two sentences p and q in which > does not occur, suppose that $[p \wedge q] = 1$ relative to K. Then (a) $p \in K$ and (b) $q \in K$. On any reasonable understanding of f, (a) entails that $f(K, p) = K$. From this and (b) we get that $q \in f(K, p)$, so by definition 4 $[p > q] = 1$ relative to K. This is to say that $p \wedge q$ is inconsistent with $\sim (p > q)$.

The fourth alternative is the possible worlds view advocated by Stalnaker. On this view, to ask whether $p > q$ is true is to ask whether q is true in a possible world that makes p is true—a p-world—and otherwise differs minimally from the actual world. More precisely, the truth conditions of $p > q$ relative to a possible world w are given in terms of a selection function f that assigns a possible world to the pair formed by p and w:

Definition 5: $[p > q] = 1$ in w if and only if $[q] = 1$ in $f(p, w)$.

Here $f(p, w)$ is understood as the most similar world to w in which p is true. From definition 5 we get that $p > q$ is true simpliciter if and only if it is true in the p-world that is most similar to the actual world.[9]

The possible worlds view does not explain fact 1. Stalnaker assumes that the p-world that is most similar to w is w itself, so that $f(p, w) = w$ when $[p] = 1$ in w. On this assumption, known as *strong centering*, definition 5 entails that if $[p] = 1$ in w, $p > q] = 1$ in w if and only if $[q] = 1$ in w. The selection function does substantive work only when the antecedent is false. Therefore, if $[p \wedge q] = 1$ in w, then $[p > q] = 1$ in w. This is to say that $p \wedge q$ is inconsistent with $\sim (p > q)$.

The foregoing considerations show that, as far as the explanation of fact 1 is concerned, definition 1 is definitely better than definitions 2-5. This is not to

[7] I owe this argument to Vincenzo Crupi.
[8] An account along these lines was initially suggested in Gärdenfors 1978 and then developed in other works such as Gärdenfors 1988, Levi 1988, and Arlo-Costa 1955.
[9] This is the view defended in Stalnaker 1991. See also Davis 1979.

say that the strict conditional view is the only view that can explain fact 1. If one adopts a variant of the possible worlds view in which $f(p, w)$ is a class of p-worlds rather than a single p-world, and the condition imposed on f is that, if $[p] = 1$ in w, then $w \in f(p, w)$—the assumption known as *weak centering*—one does not get that $p \wedge q$ is inconsistent with $\sim (p > q)$. But the point remains that the strict conditional view has a clear advantage over the four theories of conditionals considered.[10]

3. Second Argument

The second argument hinges on the following observation:

Fact 2: It seems that in some cases one can assert $\sim (p > q)$ but not $p > \sim q$.

To assert $\sim (p > q)$ is to deny that a certain relation holds between p and q. But this is not quite the same thing as to affirm that that relation holds between p and $\sim q$. So it seems that, although in many cases $\sim (p > q)$ and $p > \sim q$ are both assertable, $\sim (p > q)$ does not entail $p > \sim q$. Imagine that a detective and his assistant investigate a murder in a mansion. The three suspects are the butler, the driver, and the gardener. The butler belongs to the house staff, while the driver and the gardener belong to the grounds staff. Once some clues are collected, it turns out that the butler has an airtight alibi. Then the assistant utters the following sentence:

(5) If a member of the grounds staff did it, then it was the driver.

In this case it is reasonable for the detective to deny (5). But his denial of (5) does not imply that if a member of the grounds staff did it, then it was the gardener. Just as there is no reason to assert (5), there is no reason to assert such conditional.[11]

Fact 2 raises another controversial issue. Some theorists of conditionals hold that $\sim (p > q)$ entails $p > \sim q$ simply because they believe that $\sim (p > q)$ *means* $p > \sim q$. For example, Adams is quite explicit on this point:

> the ordinary meaning of the denial 'It is not the case that if p then q' is just to assert 'if p then not q'.[12]

However, fact 2 can hardly be dismissed by appealing to the meaning of $>$, given that the whole debate on conditionals stems precisely from the fact that it is not obvious what $>$ means. Presumably, if one takes fact 2 at face value, it is because one thinks that the assertability of a conditional requires that a certain relation obtains between its antecedent and its consequent, so that $p > q$ and $p > \sim q$ can both be denied. This divergence has direct implications on the relation between *excluded middle* (EM) and *conditional excluded middle* (CEM). If one assumes that $\sim (p > q)$ entails $p > \sim q$, one will expect that $(p > q) \vee \sim (p > q)$ entails $(p > q) \vee$

[10] This variant of the possible worlds view is in line with the account of counterfactuals sketched in Lewis 1973: 26-31, although that account is not intended to apply to conditionals. Nolan 2003 defends a version of the possible worlds view that seems to go in this direction, as it takes fact 1 into account.

[11] This example is drawn from Gillies 2004: 589.

[12] Adams 1965: 181.

($p > \sim q$). Instead, if one does not make that assumption, one will regard CEM as more dubious than EM.

Fact 1 and fact 2 are related. There is a straightforward argument to the effect that, if $p \wedge q$ is consistent with $\sim (p > q)$, then $\sim (p > q)$ does not entail $p > \sim q$. Let us call * the assumption that $\sim (p > q)$ entails $p > \sim q$. If * holds, then one can derive $\sim\sim (p > q)$ from $p \wedge q$ by CE, *modus ponens* (MP) and *reductio ad absurdum* (RAA):

1	[1] $p \wedge q$	A
1	[2] p	CE1
1	[3] q	CE1
3	[4] $\sim (p > q)$	A
3	[5] $p > \sim q$	*4
1,3	[6] $\sim q$	MP 5,2
1	[7] $\sim\sim (p > q)$	RAA 4,3,6

So it turns out that $p \wedge q$ is inconsistent with $\sim (p > q)$ because it entails $\sim\sim (p > q)$. The converse conditional, instead, is harder to justify: it is not obvious that, if $\sim (p > q)$ does not entail $p > \sim q$, then $p \wedge q$ is consistent with $\sim (p > q)$. From the premise that $p \wedge q$ is inconsistent with $\sim (p > q)$ one cannot draw the conclusion that $\sim (p > q)$ entails $p > \sim q$. Certainly, that conclusion can be obtained by means of *conditional proof* (CP):

1	[1] $\sim (p > q)$	A
2	[2] p	A
3	[3] q	A
3	[4] $p > q$	CP 2,3
1	[5] $\sim q$	RAA 3,4,1
1	[6] $p > \sim q$	CP 2,5

But not all theories of conditionals validate CP, so here it cannot be taken for granted that CP holds for >. Therefore, facts 1 and 2 deserve separate consideration: while there is a direct route from fact 1 to fact 2, there is no such route from fact 2 to fact 1.

On the assumption that fact 2 is a datum that deserves an explanation, it turns out that the strict conditional view is more credible than other theories of conditionals. The strict conditional view explains fact 2 as follows: in some cases it seems that one can assert $\sim (p > q)$ but not $p > \sim q$ because in those cases the former is true but the latter is false. It may happen that $[\sim (p > q)] = 1$ in w, in that there is a w' such that $[p] = 1$ in w' and $[q] = 0$ in w', but that $[p > \sim q] = 0$ in w, in that there is a w'' such that $[p] = 1$ in w'' and $[q] = 1$ in w''. For example, in the case of the detective one can say that (5) is false, because there are possible worlds in which its antecedent is true and its consequent is false, but that the same goes for the conditional obtained from (5) by adding 'not' in the consequent.

The truth-functional view, instead, does not explain fact 2. According to definition 2, if $[\sim (p > q)] = 1$, then $[p] = 1$ and $[\sim q] = 1$, so $[p > \sim q] = 1$. This means that, as long as assertability conditions are understood as truth conditions, $p > \sim q$ is assertable whenever $\sim (p > q)$ is assertable. Of course, one may distinguish assertability from truth and claim that fact 2 is to be explained in terms of assertability. But again, the explanatory problem then moves from the truth-functional view to the favoured account of assertability.

The probabilistic view is also unable to explain fact 2. As it turns out from section 2, on this view $[\sim (p > q)] = [p > \sim q]$. Therefore, $\sim (p > q)$ entails $p > \sim q$. This holds no matter which of the two versions of the view one choses.

Similar considerations hold for the possible worlds view. Suppose that $[\sim (p > q)] = 1$ in w. Then $[p > q] = 0$ in w, which means that $[q] = 0$ in $f(p, w)$. It follows that $[\sim q] = 1$ in $f(p, w)$, so that $[p > \sim q] = 1$ in w. This is to say that $\sim (p > q)$ entails $p > \sim q$, contrary to fact 2.

The belief revision view does not have this problem. The assumption that $[\sim (p > q)] = 1$ relative to a belief state K does not entail that $[p > \sim q] = 1$ relative to K. Since belief sets are not required to be maximal, it may be the case that neither q nor $\sim q$ belong to $f(K, p)$. By definition 4, this means that $[p > q] = 0$ and $[p > \sim q] = 0$ relative to K. So it may be the case that $\sim (p > q)$ is assertable while $p > \sim q$ is not.

Thus, the strict conditional view is not the only view that can explain fact 2. Nonetheless, as far as fact 2 is concerned, definition 1 is better than definitions 2, 3, and 5. Therefore, insofar as one thinks that fact 2 deserves an explanation, and is unwilling to take for granted that EM and CEM are equivalent, one will find that the strict conditional view has some virtue.

4. Third Argument

The third argument hinges on the following observation:

Fact 3: In some cases, $\sim (p > q)$ is paraphrased as 'It is possible that p but not q'.

This is simply a fact about ordinary language. The negation of a conditional is often expressed by using modal vocabulary. For example, the detective could easily reply to his assistant "No, it might be the gardener". Another example is the following. A philosophy teacher gives a lecture on theological matters and mentions benevolence among the properties traditionally ascribed to God. One of the students is puzzled by the very concept of divine benevolence, and utters the following sentence:

(6) If God exists, then the prayers of evil men will be answered.

In this case it is natural for the teacher to deny (6), and explain that the hypothesis that God exists does not entail that the prayers of evil men will be answered. In other words, the existence of God is consistent with the possibility that the prayers of evil men will not be answered.[13]

The issue that arises in connection with fact 3 is how can we make sense of the use of modal expressions in cases such as those just described. As is well known, modal expressions can be construed in more than one way. So it would be patently unreasonable to require that the word 'possible' in the modal paraphrase of $\sim (p > q)$ is read literally. What will be assumed here is rather that an account of conditionals explains fact 3 as long as it explains the apparent correctness of the modal paraphrase of $\sim (p > q)$ in terms of the notions it employs. As we shall see, this leaves room for a meaningful distinction between explaining and not explaining fact 3.

The strict conditional view provides the most obvious explanation of fact 3: $\sim (p > q)$ may be paraphrased as 'It is possible that p but not q' because it means

[13] This example is adapted from Stevenson 1970: 28.

precisely that it is possible that p but not q. So it makes perfect sense for the philosophy teacher to deny (6) by affirming the possibility that its antecedent is true and its consequent is false.

The truth-functional view, instead, does not explain fact 3. Definition 2 entails that it is wrong to paraphrase $\sim (p > q)$ as 'It is possible that p but not q', for it makes $\sim (p > q)$ equivalent to $p \wedge \sim q$. The implausibility of this equivalence is hardly deniable. For example, the philosophy teacher does not want to say that God actually exists.[14]

The first version of the probability view is also unable to explain fact 3. If assertability is defined in the first of the two ways suggested in section 2, then it is wrong to paraphrase $\sim (p > q)$ as 'It is possible that p but not q'. Certainly, if $\sim (p > q)$ is assertable, because $[p > q] \leq 0.5$, then it is possible that p and not q. But even if $\sim (p > q)$ is not assertable, because $[p > q] > 0.5$, it is still possible that p and not q. Since 'It is possible that p and not q' is consistent both with $\sim (p > q)$ and with $p > q$, it cannot be equivalent to the former. The point is that a conditional is often denied in virtue of the mere possibility—no matter how probable—that its consequent does not hold on the supposition that its antecedent holds. Even if it were highly probable that the driver is the murderer, the detective could still deny (5) in virtue of the possibility that the gardener is the murderer. Similarly, when the philosophy teacher corrects the student, she clearly does not intend to say that, on the supposition that God exists, is unlikely that the prayers of evil men will be answered.

The second version of the probability view improves the first in that it explains why $p > q$ may be rejected even if $P(\mathrm{q}|\mathrm{p}) > 0.5$. The case in which the detective rejects (5) although it is highly probable that the driver is the murderer can be described as one in which the value of (5) is still below the threshold fixed by the context: if $P(\mathrm{q}|\mathrm{p}) = 0.8$ but $n = 0.9$, $p > q$ is not assertable. More drastically, the case in which the philosophy teacher rejects (6) can be described as one in which the threshold fixed by the context is 1, so no lower value will do. On the other hand, this is only half of the story. We have seen that the cases in which $p > q$ can be denied are plausibly described as cases in which $\sim (p > q)$ can be asserted, and this is not what we get. Insofar as (N) holds, the second version implies that the non-assertability of $p > q$ does not amount to the assertability of $\sim (p > q)$. For example, if $n = 0.9$ and $P(\mathrm{q}|\mathrm{p}) = 0.8$, $p > q$ is not assertable, but the same goes for $\sim (p > q)$, given that it gets 0.2.

The belief revision view is no better. Even assuming that possibilities are understood as belief states, it is hard to see how definition 4 can account for the modal paraphrase of $\sim (p > q)$, given that the evaluation of $p > q$ depends on what happens in a single belief state, $f(\mathrm{K}, p)$, so the same goes for $\sim (p > q)$. In other words, 'It is possible that p but not q' cannot be read as 'There is a belief state in which p and not q', for the view implies that $\sim (p > q)$ is assertable when the unique belief state obtained from the revision of our original state is such that p and not q.

A similar problem affects the possible worlds view. As long as it is assumed that the evaluation of $p > q$ in w depends on whether q is true in a single possible world, $f(p, w)$, what is said by uttering $\sim (p > q)$ in w is that q is false in that

[14] Stevenson 1970 uses this example to show that the truth-functional view is unable to handle negated conditionals.

world. Therefore, $\sim (p > q)$ is not correctly paraphrased as 'It is possible that p but not q', for the latter sentence means that there is at least one possible world in which p and not q.

Again, this is not to say that the strict conditional view is the only view that can explain fact 3. The problem that affects the possible worlds view disappears if one adopts the variant considered in section 2, because in that case $p > q$ turns out to be false when q is false in at least one of the worlds that belong to $f(p, w)$. Nonetheless, as far as fact 3 is concerned, definition 1 is better than definitions 2-5. So the strict conditional view has an advantage over the four theories of conditionals considered.

5. Invalid Argument Forms

The arguments outlined in sections 2-4 provide some reasons for thinking that a conditional is true just in case it is impossible that its antecedent is true and its consequent is false. Although these arguments are not conclusive, given that facts 1-3 might be questioned, they suggest that there may be something right in the strict conditional view. The rest of the paper discusses two foreseeable objections to the strict conditional view, in order to show that they are not compelling.

The first objection, which goes back to Adams, concerns the examples of apparently invalid inferences that are typically used to show that nonmaterial conditionals are nonmaterial. In his paper *The Logic of Conditionals*, Adams offers nine arguments as initial evidence against the truth-functional view:

A1
(7) John will arrive at 10
---
(8) If John does not arrive at 10, he will arrive at 11

A2
(7) John will arrive at 10
---
(9) If John misses his plane in New York, he will arrive at 10

A3
(10) If Brown wins the election, Smith will retire
---
(11) If Smith dies before the election and Brown wins it, Smith will retire

A4
(10) If Brown wins the election, Smith will retire
(12) If Smith dies before the election, Brown will win it
---
(13) If Smith dies before the election, he will retire

A5
(10) If Brown wins the election, Smith will retire
(14) If Brown wins the election, Smith will not retire
---
(15) Brown will not win

A6
(16) Either Dr. A or Dr. B will attend the patient
(17) Dr. B will not attend the patient
---
(18) If Dr. A does not attend the patient, Dr. B will

A7
(19) It is not the case that if John passes history, he will graduate

(20) John will pass history

A8

(21) If you throw switches S and T, it will start

(22) Either if you throw switch S it will start or if you throw switch T it will start

A9

(23) If John will graduate only if he passes history, then he won't graduate

(24) If John passes history, he won't graduate

Adams's point is that A1-A9 are apparently invalid. Since the truth-functional view entails that A1-A9 instantiate valid argument forms, he takes this to show that the truth-functional view is seriously limited. A further claim he makes, which concerns us here, is that A1-A9 speak against the strict conditional view as well. When he comments on A4, which instantiates the principle of hypothetical syllogism, he says:

> A closely related principle is taken as a postulate in C.I. Lewis' theory of strict implication. It is unlikely, therefore, that fallacies of the kind given here can be entirely avoided by going over to formal analysis in terms of strict implication or related systems.[15]

Here Adams seems to claim that if A4 were formalized in accordance with the strict conditional view, it would be described as an instance of a valid argument form, that is:

F4

$\Box(p \supset q)$

$\Box(r \supset p)$

$\Box(r \supset q)$

So the problem of explaining its apparent invalidity would still be there. Similar considerations might be applied to some of the other examples provided by Adams. In particular, it might be contended that, if A3 and A5 were formalized in accordance with the strict conditional view, they would be described as instances of valid argument forms:

F3

$\Box(p \supset q)$

$\Box((p \wedge r) \supset q)$

F5

$\Box(p \supset q)$

$\Box(p \supset \sim q)$

$\sim p$

More generally, the objection may be phrased as follows: some arguments involving nonmaterial conditionals are apparently invalid, but according to the strict conditional view they instantiate valid argument forms; so there is something wrong with the strict conditional view.

The flaw of this objection lies in the assumption that A3-A5 instantiate valid argument forms according to the strict conditional view. Definition 1 does not

[15] Adams 1965: 168-69.

fix a unique method of formalization, that is, it does not determine a unique way to assign formulas of a modal language to conditionals. So it is not obvious that A3-A5 are to be formalized in the way suggested by Adams. In particular, as we shall see, it is consistent with the strict conditional view to claim that A3-A5 instantiate invalid argument forms that differ from F3-F5.

Let us start from the very notion of adequate formalization. One way to understand adequate formalization, which is consistent with the strict conditional view, is to assume that a formula adequately formalizes a sentence if and only if it provides a logically perspicuous representation of what is said by uttering the sentence. Since an adequate formalization of a sentence shows its logical form, this is to assume that sentences have logical form in virtue of the content they express. So the implied notion of logical form significantly differs from a syntactically oriented notion such as that adopted, among others, by Lycan, Gillies, and Kratzer. Even though a syntactically oriented notion of logical form is perfectly respectable in that it suits certain theoretical purposes, its suitability for those purposes does not prevent other notions of logical form from being equally respectable for other reasons. All that is needed here is that there is at least one coherent notion of logical form according to which logical form depends on what is said.[16]

On this understanding of adequate formalization, there are essentially two ways to formalize conditionals in accordance with definition 1, depending on how contextual restrictions on possibility are expressed at the level of logical form. One option is to incorporate such restrictions in the antecedent: $p > q$ as uttered in a context c may be represented as $\Box (p_c \supset q)$, where pc is a formula that differs from p in that it expresses a stricter condition fixed by c. That is, p_c may be read as 'p and things are like in the actual world according to c'. The other option is to incorporate such restrictions in the necessity operator: $p > q$ as uttered in c may be represented as $\Box_c (p \supset q)$, where $\Box_c$ expresses quantification over a restricted set of worlds. That is, $\Box_c$ is to be read as 'in every world that is similar to the actual world relative to c'. This second option is often associated with the idea that contexts are sets of worlds that vary as a function of the antecedent, so that any difference in the antecedent determines a difference in the intended context.[17]

No matter which of these two options one chooses, one can deny that A3-A5 instantiate F3-F5. Consider the first option. According to this option, the logical form of $p > q$ as uttered in c is $\Box (r \supset q)$, where r stands for 'p and things are relevantly like in the actual world' as understood in c. This entails that A3-A5 are adequately formalized as follows:

F3′

$$\frac{\Box (p \supset q)}{\Box (r \supset q)}$$

[16] Lycan 2001, Gillies 2009, Kratzer 2012. Iacona 2018 distinguishes a notion of logical form based on content from a syntactically oriented notion of logical form, and illustrates some significant implications of the distinction

[17] In the case of counterfactuals, the first option has been developed in different ways in in Åqvist 1973 and in Iacona 2015, while the second has been considered in various works, such as Lowe 1990. The idea that contexts vary as a function of the antecedent has been developed in Von Fintel 2001, Gillies 2007, Warmbrod 1981, and Lowe 1995.

F4′
$\Box(p \supset q)$
$\Box(r \supset s)$
$\Box(r \supset q)$

F5′
$\Box(p \supset q)$
$\Box(p \supset \sim q)$
$\sim r$

A3 is adequately formalized as F3′, rather than as F3, because the real antecedent of (10) is 'Brown wins the elections and things are relevantly like in the actual world', so the real antecedent of (11) is not a conjunction that includes the real antecedent of (10) as a conjunct. The real antecedent of (11) is rather 'Smith dies before the election, Brown wins it, and things are relevantly like in the actual world'. Similar considerations hold for A4 and A5. On the assumption that the logical form of (10)-(15) is represented in the way suggested, it turns out that A4 and A5 are adequately formalized as F4′ and F5′ rather than as F4 and F5. Since F3′-F5′ are invalid argument forms, the apparent invalidity of A3-A5 causes no trouble.

Now consider the second option. According to this option, the logical form of $p > q$ as uttered in c is $\Box_c(p \supset q)$, where $\Box_c$ is to be read as 'in every world that is similar to the actual world relative to c'. Assuming that different antecedents require different contexts, A3 and A4 are adequately formalized as follows:

F3″
$\Box_c(p \supset q)$
$\Box_{c'}((p \wedge r) \supset q)$

F4″
$\Box_c(p \supset q)$
$\Box_{c'}(r \supset p)$
$\Box_{c'}(r \supset q)$

The case of A5 is slightly different. In this case the assumption that different antecedents require different context plays no role, because (10) and (14) have the same antecedent. Nonetheless, it is plausible to postulate a context shift to explain how (10) and (14) can both be true in spite of the fact that they have contradictory consequents. So A5 is adequately formalized as follows:

F5″
$\Box_c(p \supset q)$
$\Box_{c'}(p \supset \sim q)$
$\sim p$

Since F3″-F5″ involve different operators, they are invalid argument forms. Or at least, this follows from the reasonable assumption that, in a modal language with multiple operators, the valid counterparts of F3-F5 require that the operator is fixed. Again, we get that the apparent invalidity of A3-A5 causes no trouble.

6. Material Conditionals

The second objection concerns material conditionals. As has been explained in section 1, there are cases in which conditionals are *prima facie* tractable as material

conditionals, that is, cases in which it is plausible to read $p > q$ as $p \supset q$. Therefore, it might be contended that the strict conditional view is no better than the truth-functional view as far as explanatory power is concerned: each of the two views can account at most for a limited range of cases.

Against this objection it might be argued that the strict conditional view can handle material conditionals. Let us consider an example in which it is plausible to read $p > q$ as $p \supset q$. Suppose that you submit a paper to a journal that announced a special issue on your favourite topic. After six months, you write an email to the editor of the journal to ask about the status of your submission, and you receive the following reply:

(25) Your paper was considered, if submitted before the deadline.

In this case it seems that the assessment of (25) depends on the way things actually are. If one knows whether your paper was actually submitted before the deadline and whether it was actually considered, one is in a position to say whether the assertion made by the editor is true or false.

One way to see the difference between this case and the other cases considered so far is to realize that it would be wrong to explain the truth or falsity of (25) in terms of quantification over a set of worlds that includes non-actual worlds relevantly similar to the actual world. Suppose that the editor of the journal is very unreliable and normally ignores most submissions, but that your submission was considered for purely accidental reasons. On any reasonable understanding of the relevance condition, the set of relevantly similar worlds includes worlds in which the paper was submitted but not considered, so (25) would turn out false. Yet it is plausible to say that (25) is accidentally true.[18]

The strict conditional view leaves room for cases of this kind, because it does not require that the domain of quantification includes worlds other than the actual world. It is consistent with definition 1 to say that there are cases in which the only relevant w' is w itself. In other terms, the cases in which it is plausible to read $p > q$ as $p \supset q$ are cases in which 'Necessarily, if p and things are relevantly like in the actual world, then q' is understood as 'Necessarily, if p and things are exactly like in the actual world, then q'. So the case of (25) can be described as one in which the only relevant world is the actual world. More generally, material conditionals can be treated as a limiting case in which the relevant set of possible worlds contains the actual world as its only member.

Note that conditionals about the future often behave as material conditionals. Consider the following:

(26) If I open this box, I will find chocolates

(27) If tomorrow is sunny, I will go to the beach.

In normal circumstances we take for granted that whether (26) is true or false depends on what is actually inside the box. The existence of a possible world in which the box is empty does not suffice to falsify (26). Similarly, in normal circumstances we take for granted that whether (27) is true or false depends on what will actually happen tomorrow. The existence of a possible world in which tomorrow is sunny but I stay at home does not suffice to falsify (27).

[18] Here it is taken for granted that the truth of the antecedent and the consequent of (25) does not suffice for the truth of (25) if a quantification on possible worlds is involved. But this is not essential to the point, as similar examples may be provided in which the antecedent is false, as it is shown by Kratzer manuscript, section 2.

This is not to say that the account of material conditionals offered by the strict conditional view is better than the account of material conditionals offered by the truth-functional view. At most, the former is as good as the latter. The point here is about explanatory power. While the truth-functional view works for material conditionals but has little to say about nonmaterial conditionals, the strict conditional view applies both to non-material conditionals—at least in some cases—and to material conditionals. Material conditionals can be treated as special cases of strict conditionals, whereas strict conditionals cannot be treated as special cases of material conditionals. In this respect the strict conditional view is definitely better than the truth-functional view.

7. Conclusion

Let us conclude with some general remarks about the logical significance of the strict conditional view. What has been said so far suggests that, at least in a considerably wide class of cases, conditionals are adequately formalized as strict conditionals. Therefore, as long as we restrict consideration to such cases, the logical properties of conditionals can be elucidated by employing the resources of modal propositional logic.

This is not to say that the strict conditional view works for every case. Even if it is granted that some nonmaterial conditionals are adequately formalized as strict conditionals, and that all material conditionals are handled in the way explained, the question remains of whether these two categories of conditionals are jointly exhaustive. What has been said so far is consistent with the possibility that some nonmaterial conditionals are not amenable to the formal treatment suggested.

Nonetheless, the strict conditional view poses an interesting challenge. Most theorists of conditionals tend to think that conditionals do not conform to classical logic: they claim that conditionals are not evaluable as true or false, that the arguments in which they occur are not appropriately assessed in terms of classical validity, that they violate classical rules of inferences, and so on. Insofar as the strict conditional view works, no such revisionary conclusion can be drawn. Even though propositional logic may not be enough in some cases, all we need to do is to go from propositional modal logic to its most familiar extension, modal propositional logic.[19]

References

Adams, E.W. 1965, "The Logic of Conditionals", *Inquiry*, 8, 166-97.

Adams, E.W. 1968, "Probability and the Logic of Conditionals", in Suppes, P. and J. Hintikka (eds.), *Aspects of Inductive Logic*, Amsterdam: North-Holland, 265-316.

Åqvist, L. 1973, "Modal Logic with Subjunctive Conditionals and Dispositional Predicates", *Journal of Philosophical Logic*, 2, 1-76.

[19] Previous versions of this paper were given as talks in Matera (2015), Leeds (2016), and Munich (2016). I learned much from the discussions with those who attended, before and after the talks. I would like to express my special gratitude in this connection to Vincenzo Crupi, John Divers, Kevin Mulligan, Gil Sagi, Robbie Williams, and Tim Williamson. I am also extremely grateful to two anonymous referees for their written comments on the text.

Arlo-Costa, H. 1955, "Epistemic Conditionals, Snakes and Stars", in Crocco, G., L. Farinas del Cerro and A. Herzig (eds.), *Conditionals. From Philosophy to Computer Science*, Oxford: Oxford University Press, 193-239.

Butcher, D. 1983, "An incompatible pair of subjunctive conditional modal axioms", *Philosophical Studies*, 44: 71-110.

Davis, W. 1979, "Indicative and Subjunctive Conditionals", *Philosophical Review*, 88, 544-64.

Edgington, D. 1986, "Do Conditionals Have Truth Conditions?", *Critica*, 18, 3-30.

Gärdenfors, P. 1978, "Conditionals and Changes of Belief", in Niiniluoto, I. and R. Tuomela (eds.), *The Logic and Epistemology of Scientific Change*, *Acta Philosophica Fennica*, 30, 381-404.

Gärdenfors, P. 1988, *Knowledge in Flux*, Cambridge, MA: MIT Press.

Gillies, A.S. 2004, "Epistemic Conditionals and Conditional Epistemics", *Noûs*, 38, 585-616.

Gillies, A.S. 2007, "Counterfactual Scorekeeping", *Linguistics and Philosophy*, 30, 329-60.

Gillies, A.S. 2009, "On Truth-Conditions for If (but Not Quite Only If)", *Philosophical Review*, 118, 325-49.

Iacona, A. 2015, "Counterfactuals as Strict Conditionals", *Disputatio*, 41, 165-91.

Iacona, A. 2018, *Logical Form*, Cham: Springer.

Jackson, F. 1979, "On Assertion and Indicative Conditionals", in Jackson, F. (ed.), *Conditionals*, Vol. 88, Oxford: Oxford University Press, 1987, 111-35.

Kratzer, A. 2012, *Modals and Conditionals*, Oxford: Oxford University Press.

Kratzer, A. manuscript, "Chasing Hook: Quantified Indicative Conditionals".

Levi, I. 1988, "Iteration of Conditionals and the Ramsey Test", *Synthese*, 76, 49-81.

Lewis, D. 1973, *Counterfactuals*, Oxford: Blackwell.

Lewis, D. 1976, "Probability of Conditionals and Conditional Probabilities", *Philosophical Review*, 85, 297-315.

Lowe, E.J. 1990, "Conditionals, Context, and Transitivity", *Analysis*, 50, 80-87.

Lowe, E.J. 1995, "The Truth about Counterfactuals", *Philosophical Quarterly*, 45, 41-59.

Lycan, W.G. 2001, *Real Conditionals*, Oxford: Oxford University Press.

McDermott, M. 2007, "True Antecedents", *Acta Analytica*, 22, 333-35.

Nolan, D. 2003, "Defending a Possible-Worlds Account of Indicative Conditionals", *Philosophical Studies*, 116, 215-69.

Stalnaker, R.C. 1970, "Probability and Conditionals", *Philosophy of Science*, 37, 64-80.

Stalnaker, R. 1991, "A Theory of Conditionals", in Jackson, F. (ed.), *Conditionals*, Oxford: Oxford University Press, 28-45.

Stevenson, C.L. 1970, "If-iculties", *Philosophy of Science*, 37, 27-49.

von Fintel, K. 2001, "Counterfactuals in a Dynamic Context", in Kenstowicz, M. (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, 123-52.

Warmbrod, K. 1981, "Counterfactuals and Substitution of Equivalent Antecedents", *Journal of Philosophical Logic*, 10, 267-89.

Sangiovanni, A., *Humanity without Dignity. Moral Equality, Respect, and Human Rights.*
Cambridge (MA): Harvard University Press, 2017, pp. x + 320.

Although one should be cautious to call a philosophical book beautiful, I think that *Humanity without Dignity* properly deserves such qualification. Not only does it display rigor and clarity in developing the complex and interrelated arguments, and not only does it advance an original thesis concerning the grounding of human rights, but it is also beautifully written and shows a mastery of classical texts and literature which is unusual in an analytical work. The result is an enrichment of the argumentation with historical depth and literary examples, which makes the reading truly enjoyable.

The book's main question is the following: human rights presuppose moral equality among humans; in turn, moral equality is usually accounted for on the basis of our dignity, which constitutes the status demanding recognition and respect. In most accounts of the grounds for moral equality and human rights, dignity is singled out as the core of human worth, the kernel of our common humanity. Sangiovanni disputes this prevalent account, carefully criticizing the three main views on dignity (ch.1), namely the Aristocratic view (from Aristotle, to Cicero, to Baldassar Castiglione), the Christian view and the Kantian view. Then he presents his alternative (ch.2), based on a negative approach to the issue, by means of analyzing and reflecting on the reactive attitudes governing the practices of treating others as inferiors. From such a reflection, it emerges that cruelty is the wrong displayed in the different ways of treating others as inferior, and cruelty is defined as the unauthorized and wrongful use of others' vulnerability to attack or obliterate people's capacity to develop and maintain an integral sense of self. Then he takes up a thorough analysis of discrimination in order to illustrate how social cruelty works in demeaning, obliterating and deleting the sense of self of discriminated persons (ch.3).

In the second part of the book, Sangiovanni turns to human rights, looking for a concept that is consistent with the previous discussion of moral equality and of the harm produced by inferiorizing treatments. In line with his argumentative approach in the first part, he defines human rights in the negative, as "those moral rights whose systematic violations ought to be of universal, legal and political *concern*"; thus, through violations engendering universal concern, human rights can be identified and asserted (ch.4). Equipped with this concept, then, he proceeds discussing some central topics in the philosophical discussion of human rights, namely the moral bases of international human rights and the distinction between basic or, better, fundamental from non-fundamental rights (ch.5). Finally, he wonders whether we have an obligation to pursue the protection offered by human rights at the international level and to embody such protection in a system of international norms (ch.6). Viewing human rights as primarily meant to protect people from attacks on their equal moral status enables him to answer many open questions, for example, which rights are fundamental and hierarchically prior to others, and adds a philosophical depth to purely legal and political approaches more focused on enforcement and on the list of the human rights we have. I think that the moral equality perspective, forcefully put forward by Sangiovanni, is indeed the key to understanding and supporting the international system of human rights, and, much more than other approaches to

Argumenta 4, 1 (2018): 193-206
ISSN 2465-2334
© 2018 University of Sassari
DOI 10.14275/2465-2334/20187.boo

global justice, such as luck egalitarianism or utilitarianism, makes sense of what is intolerable in certain circumstances of life beyond poverty and deprivation.

I shall focus my critical discussion on two points of the first part of the book concerning which I have some reservations, namely the notion of the integral sense of self—and its role in the architecture of the main argument—and the view of respect as opacity respect. Let us start by considering Sangiovanni's negative approach to moral equality. Instead of looking for the common property in which human worth consists, Sangiovanni, provisionally assuming moral equality, examines its violations and wonders what is wrong about them. All major forms of treating people as inferior (stigmatizing, dehumanizing, infantilizing, objectifying, instrumentalizing) share social cruelty as their common denominator, and what defines cruelty is not just the harm and the injuries produced, but the demeaning attitude for cruelty aims at attacking or destroying the integral sense of self, taking advantage of others' vulnerability. Thus, it is the integral sense of self the fundamental good and crucial interest shared by all human beings, beyond their different capacities, circumstances and projects. This notion, which is reached through the negative approach, is able to satisfy the two desiderata which the grounding of moral equality should respond to (and which dignity fails to satisfy), insofar as it explains a) why we are morally equal (equality desideratum) and b) why moral equality is worthy and should be protected by rights (rationale desideratum). The alternative to dignity is therefore not another property, supposedly, common to all human beings, but it is rather the central human interest to develop and preserve a sense of self. The grounds of moral equality consist not in the kernel of human value shining inside any human being, but rather in what we most care about, which makes us all vulnerable to wrongful violations and hence requires protection via moral rights. The rejection of social cruelty, implicit in all inferiorizing treatments threatening the integral sense of self, calls for respect and moral rights. Sangiovanni thinks he has disposed in this way of the main difficulty concerning the possession of the property which makes us *digni*, worthy of equal consideration and respect, namely the actual variation in rational capacity and rational deliberations from which some human beings (small children, severely disabled individuals, victims of Alzheimer and senile dementia) are in fact excluded.

Moral equality requires that the reciprocal relationship within the moral and social community be governed by respect. The kind of respect relevant for Sangiovanni is "recognition-respect", according to a well-known distinction made by Stephen Darwall,[1] that is, the respect that we owe each other unconditionally, just as (equally vulnerable) human beings, and not the "appraisal-respect" which is attributed on the basis of achievements and merits and is not equal. Moreover, the recognition-respect here in order is also "opacity-respect"[2] for it implies restraint confronting others, keeping the right distance to protect the self-presentation of other people without exposing them to inquisitiveness, rudeness and discomfort.

Generally speaking, the negative approach used by Sangiovanni has clear advantages over alternatives when dealing with human values, a highly sensitive area for disagreement. It is indeed easier to find agreement on the intolerability of the violation of a given value, and the reactive attitudes to violation provide

[1] Darwall, S. 1977, "Two Kinds of Respect", *Ethics*, 88, 39-49.

[2] Carter, I. 2011, "Respect and the Basis of Equality", *Ethics*, 121, 538-71.

insight to single out a special value on which we could agree by implication. Yet, here is precisely my critical point: why is it that the value or fundamental good that the intolerable violation to our moral equality points to is the integral sense of self, instead of dignity? I understand that dignity cannot be the starting point for the argument in favor of equal human right, for dignity, literally, means that human beings are worthy (*digni*) but it requires an account of a) why we are worthy, in virtue of what property, and b) how come we are all equally worthy. Sangiovanni believes that none of the responses in the three traditions of dignity is satisfactory; hence, he takes the different negative route to moral equality. In this way, he is able to establish that inferiorizing treatments, taking advantage of human vulnerability, are socially cruel, hence moral equality ought to be presupposed to make sense of our reactive attitudes. Yet, at this point, why is it that the inferiorizing treatment attacks the integrity of our sense of self, instead of our dignity? It seems to me that the greatest good of the integral sense of self does not necessarily follow from the intolerability of social cruelty and of inferiorizing treatments. Why cannot the sense of self be the sense of one's worth, hence of one's dignity? I conjecture that his reason to favor the sense of self over dignity lies in two issues linked to dignity above mentioned. The first issue is to specify what human worth consists in, what is the special human value in virtue of which all human beings have dignity, and the answer is usually found in rational capacities, both in the Christian and in the Kantian tradition, though differently specified. The second issue is precisely connected with this answer, for, first, the rational capacities are not present in all human beings equally, and, second, in some of them, like in the severely mentally handicapped or in very small children, are absent. Thus, it seems that dignity cannot be the ground for moral equality. I think however that similar issues can be raised concerning the integral sense of self. Sangiovanni's argument is based on the difference between a property and a crucial center of our care and concern. The latter does not presuppose equal intellectual and moral capacities, and everyone, no matter how clever and morally accomplished, cares about oneself. Yet if the care for oneself is understood as the instinct of self-preservation, this is certainly something that we all share, but also something that goes beyond humanity, encompassing all living beings. Sangiovanni, however, means something more distinctly human, namely the capacity of seeing oneself, and of constructing and reconstructing one's image according to what one wants to be, as well as the capacity to present oneself to others so as to be socially recognized. Another part of an integral sense of self is the gap between how we see ourselves and how we want to be seen, which is often a reason for self-improvement, as well as the reason to limit our social exposure. Clearly though, caring for the integral sense of self implies the *capacity* of developing, revising and reshaping our images according to our commitments and wishes. And not all human beings share this capacity in the same measure across the board and some people are completely deprived of a proper sense of self. Thus, it seems to me that the issues connected with grounding moral equality on dignity, in a way resurface here. For no matter what the grounds for moral equality, the problems a) of human variations and b) of human beings that are not autonomous persons in the proper sense arise and cannot easily be disposed of, even adopting a negative approach.

The problem of human variations, whether concerning the capacity of rational agency or of having an integral sense of self, has been, to my mind, satisfactorily responded to with reference to the range property, that is a non-variant

property supervening over a range of possible variations of the variant property.[3] Sangiovanni, too hastily, dismisses the range property argument because he says that it is still to be explained why the equal possession of the range property should count more than the unequal possession of the underlying property in the highest degree. I think that the answer can be found in Sangiovanni own's argument. He has explained that the way to get to the basis of equality, i.e. the sense of self—expressed in a range property following my suggestion—is through our reactive responses to the various forms of inferiorizing treatments. From there we arrive at the intolerability of the social cruelty underlying such kinds of treatment, and to the crucial importance of the sense of self. The responsive attitudes to violations are actually independent of how deep and articulated is the sense of self, and of whether it is the product of autonomous reflection or induced by social conditioning. It is from outside that we react to the violations, and how well developed is anyone's sense of self does not count in judging the violation intolerable. The sense of self is ascribed from outside, hence it is a range property, which we presume present in everyone and which makes us indignant at its violations. Consider now the second issue of moral equality, however grounded, namely the fact that some human beings are not autonomous and seem deprived of the capacity of having a sense of self as well (or the rational and moral capacities) above a certain threshold. Here, again, I think that Sangiovanni's negative approach can be helpful: if the starting point is the violations and our consequent reactive attitudes, the latter are even stronger the more vulnerable is the victim. This establishes the moral status of the victim, no matter if deprived of an integral sense of self, for the victim is the recipient of our reactive attitudes, of our care and affection, and capable of reciprocating our affection and love. Such moral status deserves protection by rights and respect by us, though rights and respect are not equal insofar as these persons are not recognized as autonomous. The negative approach adopted by Sangiovanni can indeed help in solving both issues, but, in my view, it does not change if the ground of moral equality is dignity rather than the integral sense of self, at least once dignity is not assumed as the first premise but as the arrival point of a negative argument proceeding from violations.

The second critical remark I should like to make concerns the view of respect as opacity respect, which Sangiovanni elaborates on the basis of Carter's.[4] If moral equality is ultimately grounded in having (developing and preserving) an integral sense of self, then respect is a kind of restraint against coming too close to people's sense of self. Such a distance is required in order to protect the sense of self from violations, from social cruelty, and a common and daily violation is being exposed to the public gaze without our consent. I do not dispute that certain kinds of unauthorized exposure are disrespectful, but is this sufficient for characterizing all there is to respect for persons as opacity and distance? Sangiovanni subscribes to Darwall's notion of recognition respect as the relevant form of respect governing relationships among equals. Yet recognition respect is always also *recognition* of the person as a *person* and as an *equal*. While respectful actions vary in different contexts, they are always accompanied by *an attitude of regard* for the other person, which precisely represents the recognition element in respect. Actually, I think that one kind of violation of the equal moral status of

[3] Carter, I. 2011, cit.
[4] Carter, I. 2011, cit.

persons, and an attack on their sense of self, aimed at inferiorizing others, is invisibility, which, curiously enough, does not figure in the list of forms of inferiorizing treatment analyzed by Sangiovanni. Is it not the case that respecting members of oppressed groups as persons, via obscuring their life and circumstances, bracketing the latter as irrelevant, implies reproducing their invisibility *qua* members of their group? Is it not a way of recognizing them as persons only *beyond* who they are and in a way dispensing with their membership in the oppressed group?[5] If invisibility is included in the list of inferiorizing treatments, the question is: can the quest for recognition, regard and consideration, especially crucial in case of historic discrimination, be reconciled with opacity respect? In a sense, Sangiovanni suggests such reconciliation when he says that, in order to respect people as persons, we have to take them as self-presenters claiming to be recognized according to their own modes of presentation (89). In that case, the opacity would concern the content of the personal presentation and perspective not to be scrutinized closely and exposed unnecessarily, while the individualizing act of recognition would concern the gaze of regard towards others. I am not sure Sangiovanni would agree with my amendments but it is worth noting that his own argument provides a basis for the solution to this problem.

University of Eastern Piedmont ANNA ELISABETTA GALEOTTI

[5] Galeotti, A.E. 2010, "Respect as Recognition. Some Political Implications", in Seymour, M. (ed.), *The Plural States of Recognition*, Basingstoke: Palgrave Macmillan, 78-97.

Iacona, A., *Logical Form: Between Logic and Natural Language*. Cham: Springer International, 2018, pp. vi + 133.

The central tenet of Andrea Iacona's book is that "two notions of logical form must be distinguished: according to one of them, logical form is a matter of syntactic structure; according to the other, logical form is a matter of truth conditions. […] In the sense of 'logical form' that matters to logic, logical form is determined by truth conditions" (v). The work is composed of three parts: chapters 1-3 provide a broad-stroked, yet informative history of the notion of logical form from Aristotle to date; chapters 4-6 constitute the core of the book, where a novel truth-conditional understanding of logical form is articulated; chapters 7-9, finally, discuss a number of applications of the core theory in such areas as logic, epistemology, and the semantics of natural language. As most of the non-historical material had previously appeared in print, this book bears witness to the author's nearly decade-long engagement with the topic.[1]

It is a familiar empirical observation that we are able to grasp the meaning of all sorts of complex sentences, even ones we have never heard before. The standard explanation for such cognitive abilities of ours is that the semantics of

[1] Iacona, A. 2010, "Validity and Interpretation", *Australasian Journal of Philosophy*, 88, 247-64; Iacona, A. 2013, "Logical Form and Truth Conditions", *Theoria*, 28, 439-57; Iacona, A. 2015, "Quantification and logical form", in Torza, A. (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Cham: Springer, 125-40; Iacona, A. 2016, "Two Notions of Logical Form", *Journal of Philosophy*, 113, 617-43; Iacona, A. 2016, "Vagueness and Quantification", *Journal of Philosophical Logic*, 45, 579-602.

natural language is compositional: the meaning of a complex sentence is specified by its structure together with the meaning of its constituents. On a view that has been developed by Tarski, Davidson and Montague among others, the meaning-defining structure of a sentence is its logical form. For example, the meaning of 'John is tall and Betty is at the party' is uniquely specified by the meaning of its atomic subsentences, together with the semantic rule associated with sentences of the form '$p \wedge q$'. This is what Iacona calls the 'semantic role' of logical form (31).

We are also familiar with the adage that an inference is valid in virtue of its form. For example, 'John is tall and Betty is at the party. Therefore, Betty is at the party' is valid because it instantiates the schema 'p $\wedge$ q. Therefore q'. According to a view, fully articulated by Frege and Russell but tracing back to ancient syllogistic, the feature of a sentence which explains its inferential properties is its logical form. This is what Iacona calls the 'logical role' of logical form (24).

According to the 'current conception', there is a unique notion of logical form playing both the semantic and the logical role (36). On that conception, the logical form of a natural language sentence, which can be exhibited in a suitable formalization (typically, carried out in first-order logic or some extension thereof), accounts for the meaning of a sentence, as well as its inferential properties. Iacona's central thesis is that the current conception is misguided, as different theoretical roles require different notions of logical form (38).

The current conception has it that logical form is an intrinsic property of sentences. For instance, the influential view developed by Montague identifies the logical form of a natural language sentence with its deep syntactic structure, that is, the syntactic structure of its formal regimentation. (Other, not merely syntactic versions of intrinsicalism about logical form are discussed in the book, such as Davidson's semantic approach.) Iacona does not take issue with intrinsic logical form's ability to play the semantic role: once a natural language sentence *s* is regimented as a formula *r*, the syntax of *r* suffices to specify the meaning of *s* as a function of the meaning of *r*'s subformulas.

The author's first key claim is that the intrinsicalist conception of logical form cannot fulfill the logical role. In some special cases, it can: the aforementioned inference 'John is tall and Betty is at the party. Therefore, Betty is at the party' can be correctly assessed as valid by simply attending to the syntactic structure of its premise and conclusion. On the other hand, if someone says 'this is not this' while pointing first at an object *a*, and then at a distinct object *b*, it is clear that the speaker has not uttered an inconsistency. If so, the sentence is correctly formalized as '$a \neq b$', although none of its intrinsic properties forces such a formalization, as opposed to the inconsistent, and incorrect 'a $\neq$ a' (48). Likewise, the inference 'Now it is raining. Therefore, now it is raining' is invalid, as one might utter the premise at a time *t*, when it is in fact raining, and the conclusion at some later *t'*, when the rain has stopped. This fact suggests that the form of the argument is '$\phi(t)$. Therefore $\phi(t')$', rather than '$\phi(t)$. Therefore $\phi(t)$', as the intrinsicalist view would suggest.

The moral that Iacona draws from the above observations is that no intrinsic property of an utterance can play the logical role. According to his *truth-conditional notion* of logical form, such a role is played instead by an extrinsic property of an utterance, namely by a property it has in virtue of the proposition it expresses relative to an interpretation (63). Propositions are here individuated

hyperintensionally, since we want to distinguish the logical form of necessarily equivalent sentences such as '1 + 1 = 2' and 'if snow is white, then snow is white'. When it comes to modeling propositions, the author seems to favor Kit Fine's theory of truthmaker content, which identifies propositions with sets of states—structureless abstract entities acting as the verifiers of statements. Accordingly, the propositional content of a sentence 'p' is the set of all its possible verifiers. The content of '1 + 1 = 2' is therefore distinct from the content of 'if snow is white, then snow is white', since the first, but not the latter includes states about numbers.[2] It is worth remarking that the truth-conditional notion of logical form can be substantiated by alternative accounts of fine-grained content (57).

Although Iacona does not develop an algorithm for producing the (truth-conditional) logical form of sentences of any particular fragment of natural language, he lays the groundwork by offering an adequacy condition (AC) for all such possible algorithms, according to which logical form is just as fine-grained as propositional content. More precisely:

> (AC) Any interpreted sentences 'p' and 'q' have the same formalization, up to equivalence,[3] if and only if they express identical propositions (58).

This condition is meant to encode the key thesis that logical form, in the sense playing the logical role, is determined by propositional content, and is therefore an extrinsic property of sentences.

As Iacona argues, the aforementioned counterexamples to the intrinsicalist conception of logical form can be handled by an adequate formalization. On the Finean view of content, the proposition expressed by an utterance of 'this is not this' in the relevant context is the set of all states that make it the case that a is not b; accordingly, the logical form of 'this is not this' is expressed by the formula 'a ≠ b'. The argument 'Now it is raining. Therefore, now it is raining' is invalid, when premise and conclusion are uttered at different times t, t'. For in such a context, the proposition expressed by the premise only includes states that make it the case that it rains at t, whereas the proposition expressed by the conclusion only includes states that make it the case that it rains at t'; accordingly, the argument exemplifies the schema '$\phi(t)$. Therefore $\phi(t')$'.

It appears, however, that the conjunction of Iacona's truth-conditional notion of logical form (as constrained by AC) and Fine's theory of truth-maker content may not be sufficiently general. For example, on Fine's theory the content of a natural language sentence '*P* and *Q*' is the set of states $s \sqcup t$ such that s and t are verifiers of '*P*' and '*Q*', respectively, and $s \sqcup t$ is their mereological fusion. Since, in general, $s \sqcup s = s$, the content of any sentence '*P*' is identical to the content of '*P* and *P*'. By AC, '*P*' and '*P* and *P*' have therefore the same logical form, up to equivalence. But arguably we had better not conflate the logical form of '*P*' and '*P* and *P*', if we want to acknowledge contexts in which it is correct to say that the logical form of the latter, but not the former is conjunctive.

[2] See Fine, K. 2017, "A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation", *Journal of Philosophical Logic*, 46, 6, 1-50; Fine, K. 2017, "A Theory of Truthmaker Content II: Subject-matter, Common Content, Remainder and Ground", *Journal of Philosophical Logic*, 46, 6, 675-702.

[3] Given a suitably fine-grained equivalence relation for formulas which is stronger than the strict biconditional.

The issue can be addressed either by weakening AC, allowing sentences with the same content to have different logical form, or by adopting a theory of propositional content more general than Fine's.

Turning to some applications of his proposals, Iacona employs the truth-conditional notion of logical form in order to shed light on Frege's puzzle: what makes it so that 'Hesperus is Phosphorus' is informative, whereas 'Hesperus is Hesperus' is not, if the two sentences express the same proposition? The author indeed agrees that those sentences express the same proposition, which is captured by the formula '$h = h$' (or equivalently, '$h = p$'). What explains their difference in epistemic status is that it is not trivial that one and the same formula adequately captures the logical form of both 'Hesperus is Phosphorus' and 'Hesperus is Hesperus', and so it is not trivial that the two sentences are logically equivalent (80). The philosophical moral is that logical knowledge (in this particular case, knowledge that '$h = h$' is a logical truth) is compatible with ignorance of logical form (such as ignorance of the fact that '$h = h$' is the logical form of 'Hesperus is Phosphorus'). Nevertheless, it is worth mentioning that the truth-conditional account does not force upon us Iacona's own solution to Frege's puzzle, as he could have just as well claimed that 'Hesperus is Phosphorus' and 'Hesperus is Hesperus' express distinct propositions (given a suitably fine-grained theory of content), and so by AC that they have non-equivalent logical forms.

Another application of Iacona's theory concerns the logic of quantifier phrases, such as 'all', 'some', 'most', 'half' etc. On the truth-conditional account, the logical properties of sentences like 'all whales are mammals' and 'half the philosophers are wise' are supposedly accounted for by their first-order formalization. A potential objection to the present approach is that not all quantifier phrases are first-order definable: although we can translate 'some As are Bs' into predicate logic, it is not possible to do the same with 'half the As are Bs'. However, Iacona reminds us that "if logical form is determined by truth conditions […], formalization is representation of content rather than translation" (106). So, although 'half' is not first-order definable, 'half the philosophers are wise' can nevertheless be formalized as $\exists_2 x(Px \wedge Wx)$ (i.e., '2 philosophers are wise') if there are 4 philosophers, as $\exists_3 x(Px \wedge Wx)$ (i.e., '3 philosophers are wise') if there are 6 philosophers, etc. The same line of reasoning carries over, *mutatis mutandis*, to a number of other first-order undefinable quantifier phrases.

I myself have misgivings about Iacona's proposed solution. For his strategy hinges on the key assumption that the content, and thus the logical form of 'half the *A*s are *B*s' is a function of the actual extension of *A*, *B*. But then, by the same token, the content, and thus the logical form of 'some *A*s are *B*s' should also be a function of the actual extension of *A*, *B*, in such a way that, if there are n wise philosophers, 'some philosophers are wise' gets formalized as $\exists_n x(Px \wedge Wx)$—by all means an overly revisionary view of the logical form of existential quantifier phrases, which Iacona does not endorse, and rightly so. This disanalogy between the logical form of first-order definable vs undefinable quantifier phrases cries out for explanation. One might address the issue by claiming that a sentence with a first-order definable quantifier phrase, such as 'some philosophers are wise', has in fact two candidate formalizations: $\exists_n x(Px \wedge Wx)$, which is a function of the actual domain, and $\exists x(Ax \wedge Bx)$, which is domain-independent—and that the latter takes priority. However, Iacona doesn't provide any princi-

pled reason for choosing the domain-independent formalization over the alternative. Moreover, if domain-independent formalizations are indeed to be preferred, a uniform way to formalize quantifier phrases presents itself. For if we assume that logical form is expressed in a second-order language, all quantifier phrases are definable, and so 'half the philosophers are wise' will be formalized by the standard domain-independent truth-condition $\frac{|P \cap W|}{|P|} = \frac{1}{2}$ (i.e., 'the number of wise philosophers is half the number of philosophers'). Notice that this alternative solution is consistent with Iacona's central tenet that logical form is determined by content, provided that the content of a quantified sentence is not a function of the actual domain of interpretation. It is also consistent with Fine's view that the content of a sentence is the set of its possible truth-makers, and not just the possible truth-makers that agree with the actual ones in matters of domain.

Logical Form touches upon a number of other applications and discussions for which there is no space here. The book covers a vast ground, and does so by meeting the highest standards of clarity and rigor, without ever getting overly technical. Moreover, most chapters are self-consistent, which allows a more focused approach to the work. I warmly encourage anyone working in the philosophy of logic, language, and in linguistics, to read *Logical Form*.

National Autonomous University of Mexico (UNAM) ALESSANDRO TORZA

Bliss, R. and Priest, G. (eds.), *Reality and its Structure: Essays in Fundamentality*.
Oxford: Oxford University Press, 2018, pp. vii + 311.

Bliss and Priest's edited volume *Reality and its Structure* collects together fifteen essays on metaphysical grounding. Unlike its nearest predecessors—Correia and Schnieder's *Metaphysical Grounding* (Cambridge University Press, 2012) and Hoeltje, Schnieder and Steinberg's *Varieties of Dependence* (Ontos Verlag, 2013)—this collection is less about introducing and defending the notion of grounding and more about challenging the received view. It is similar, in this respect, to Mark Jago's *Reality Making* (Oxford University Press, 2016), even though it appears to be much more focused than this latter.

In their introductory essay ("The Geography of Fundamentality") Bliss and Priest argue that the received view of metaphysical grounding amounts to the conjunction of four theses:

1. *The hierarchy thesis*: Reality is hierarchically structured by metaphysical dependence relations that are *anti-reflexive*, *anti-symmetric*, and *transitive*;
2. *The fundamentality thesis*: There is some thing(s) which is fundamental (that is, such that there is nothing on which it depends);
3. *The contingency thesis*: Whatever is fundamental is merely contingently existent;
4. *The consistency thesis*: The dependence structure has consistent structural properties.

Bliss and Priest offer a useful taxonomy of the alternatives to the received view (7-13) based on the combination of four structural properties: *anti-reflexivity* (AR), *anti-symmetry* (AS), *transitivity* (T), and *extendability* (E), viz. the claim that

everything metaphysically depends on something else. The *hierarchy thesis* corresponds to the combination of *AR*, *AS*, and T, while the *fundamentality thesis* corresponds to ¬*E*. Since ¬*AR* implies ¬*AS*, and ¬*AS* and *T* imply ¬*AR*, there are *five* possible combinations of the first three properties and their negations. In particular, if *AR* or *T* are put aside, *circles of ground* become possible. Moreover, all these five combinations are compatible both with *E* and with its negation. In particular, the received view is a form of *metaphysical foundationalism* (F), according to which each dependence chain *terminates in a foundational element*; notable alternatives to foundationalism are *infinitism* (I), according to which there are no foundational elements, and *coherentism* (C), the most radical version of which states that everything metaphysically depends on anything else.

The main question that this volume tackles is whether there are compelling reasons for taking foundationalism to be the *right* picture of reality. Before moving on, we must get clear of some misconceptions. First, some might suggest that foundationalism is true *by definition*, since the notion of grounding *is defined* as being anti-reflexive, anti-symmetric, transitive, etc. A related point would be that that notion was introduced *at some theoretical purpose*—for example, in order to deliver a certain picture of reality—and it can serve this purpose only if it has all the properties listed above. Of course, we can define the notion of grounding just in this way; the question is not, however, about our definitions, but about the reasons for which we believe that reality *responds* to those definitions (cf. Bliss' essay, 74). Even so, one might argue that the received view does not need *arguments*, since it is *obvious* that grounding is anti-reflexive, anti-symmetric, transitive, etc.; we agree with the Editors that these intuitions are either empty or not trustworthy (10). The *arguments* for the received view divide into two categories. On the one hand, some arguments are *metaphysical*: they offer reasons to believe that the world cannot be such as to contain metaphysical dependence relations which violate the standard view; on the other hand, other arguments are *explanatory*: since grounding is intimately connected with metaphysical explanation, it may be that it cannot violate the standard view without losing such connection. As the essays in this collection make abundantly clear, however, the alternatives to foundationalism have been grossly underestimated in the current literature, and, for this reason, foundationalism itself is poorly defended—if defended at all.

The volume is divided in three parts: Part I addresses the *hierarchy thesis*, Part II addresses the *fundamentality thesis*, and Part III is about the *contingency* and the *consistency theses*. The first two parts are more substantial (7 and 6 essays respectively), while the third one is significantly shorter than the other two (2 essays). I will proceed by summarizing the main contributions of each essay (leaving some comments in brackets).

1. The Hierarchy Thesis

Gabriel Rabin ("Grounding Orthodoxy and the Layered Conception") explores the connections between the received view and the *layered conception of reality*, that is, the idea that dependence relations structure reality into a hierarchy of levels. Rabin argues that non-standard conceptions of grounding, which give up on one or more of its structural properties, are still able to recover this layered conception; in particular, the layered conception is compatible both with *non-irreflexivity* (grounding is not always reflexive) and *reflexivity*, and with *non-*

asymmetry (grounding is not always asymmetric) even if not with *symmetry*. Moreover, the layered conception is compatible with failures of transitivity. Appealing to the fact that the notion of metaphysical dependence has been introduced *with the purpose of* delivering a layered picture of reality offers therefore little advantage to metaphysical foundationalism over many of its alternatives.

Elizabeth Barnes ("Symmetric Dependence") claims that metaphysical dependence must be understood as *non-asymmetric* rather than anti-symmetric. Her examples include, among others, the following one. Arguably, the evacuation of Dunkirk is an essential part of World War II: WWII wouldn't have been the same if that evacuation had not occurred. At the same time, what the evacuation of Dunkirk is depends, at least in part, on its being part of WWII: if that evacuation wasn't occurred as part of WWII, it wouldn't have been *the evacuation of Dunkirk*. So, it does seem that both the evacuation of Dunkirk depends on WWII, and that WWII depends on the evacuation of Dunkirk.

Ricki Bliss ("Grounding and Reflexivity") focuses on *circles of (immediate) ground*, that is, circles which are formed by *reflexive* instances of metaphysical dependence. She claims that the most compelling arguments against circles of ground are *explanatory*—self-grounded entities would give rise either to *viciously circular* metaphysical explanations, or to *explanation failures*—rather than *metaphysical*—self-grounded entities would, in some sense, 'bootstrap themselves into being'. In her careful examination of these arguments, Bliss points out that metaphysical foundationalism is not obviously better off, from the explanatory point of view, than a theory which posits circles of ground: even if self-grounded entities metaphysically explain their own, foundationalism routinely posits entities for whose existence there is no explanation at all.

Daniel Nolan ("Cosmic Loops") considers the possibility of loops which go around all the levels of reality. An example of a 'cosmic' loop is an *Aleph-world*, in which there is an object—the *Aleph*—which contains anything else as a proper part, including the Aleph itself (the example is taken from Borges' famous short piece). An Aleph-world requires giving up either on *AR* or *T*. Nolan explores this second path: he suggests that, even if cosmic loops require dependence relations to be non-transitive, these relations could still be *locally* transitive, that is, transitive in a sufficient small portion of the cosmic loop. (Assume that this strategy could be generalized—that is, suppose that grounding is non-irreflexive even if it is *locally* anti-reflexive, non-asymmetric even if it is *locally* anti-symmetric, etc.. This would have some notable consequences for the epistemology of *grounding*: we could have *justified metaphysical beliefs* about dependence's having certain properties even if these properties fail in the periphery).

Thompson ("Metaphysical Interdependence, Epistemic Coherentism, and Holistic Explanation") argues for a form of *metaphysical interdependence* (MI), that is, the view that *(i)* there are no foundational facts and *(ii)* one entity can appears in its own metaphysical ancestry. MI requires to give up both on *AS* and on *E*. Thompson supports MI with an analogy with epistemic coherentism; Moreover, she points out that the friends of metaphysical interdependence stand in need to recognize *holistic metaphysical explanations*, that is, metaphysical explanations of an entity in terms of a cluster of other entities which may depend on that entity itself (further support to holistic metaphysical explanation is provided by Barnes, 65-7).

Graham Priest ("Buddhist Dependence") connects the contemporary debate with Eastern philosophical traditions. Metaphysical dependence is particularly central to Buddhist philosophy. Priest points out that, in the context of these traditions, foundationalism stands beside coherentism, and the standard conception of grounding (from the point of view of the contemporary debate) stands beside non-standard ones, according to which grounding is *non-well-founded*. (Notice that many of the views considered by Priest are about what *conceptually* depends on what; another debate which may highly profit from being informed by Eastern philosophies is the one on *conceptual grounding*).

Jon Litland ("Bicolletive Grounds") formulates a new account of *bicollective* ground. Grounding is said to be *left-collective* if there are some truths $\gamma 1$, $\gamma 2$, $\gamma 3$, etc. such that $\gamma 1$, $\gamma 2$, $\gamma 3$ ground ϕ *taken together*, without any of them grounding ϕ on its own. Grounding is said to be *right-collective* if there is some truth ϕ such that it grounds $\gamma 1$, $\gamma 2$, $\gamma 3$, etc. *taken together*, without ϕ grounding any of these truths. Finally, grounding is said to be *bicollective* if it is both right- and left-collective. Bicollective ground can be used to formulate, in particular, a form of *coherentism without circles,* since it might be that $\gamma 1$, $\gamma 2$, $\gamma 3$ depend on ϕ and ϕ depends on $\gamma 1$, $\gamma 2$, $\gamma 3$, without either $\gamma 1$ depending on ϕ neither ϕ depending on $\gamma 1$.

2. The Fundamentality Thesis

Einar Duenger Bohn ("Indefinitely Descending Grounds") argues that dependence relations could not be—and even actually aren't—well-founded. Basically, his argument is that *hunky* worlds—that is, worlds which are both *gunky*—viz., such that everything in those worlds has a proper part—and *junky*—viz., worlds in which everything is a proper part of something else—requires that dependence relations are not well-founded; he then claims that we have inductive reasons to think that our world is actually hunky.

Notice that the main argument for taking dependence relations to be well-founded comes from considerations about *metaphysical explanation*: non-well-founded dependence relations would give raise to *vicious regresses* inasmuch as a complete—or *completely satisfactory*—explanation of any entity could not be provided. Bliss and Priest (18-9) point out that these argument can also be formulated in a way that fails to make a direct appeal to vicious regresses: the point would rather be that *non-fundamental* entities wouldn't be apt to provide the kind of complete metaphysical explanation which is the foundationalist's target. More in general, there is much to be learnt from this book about grounding and explanation. A road-map will be useful to the interested reader. Bliss (84-8) focuses on the role of explanatory argument in motivating *anti-reflexivity*. Barnes (55-60) argues that metaphysical explanations can be symmetric. Finally, Barnes (65-7), Thomson (119-23), and Litland (143-4) argue that metaphysical explanations can be *holistic*.

Kelly Trogdon ("Inheritance Arguments for Fundamentality") questions the cogency of *reality inheritance arguments* for foundationalism. An influential version of this argument (basically due to Jonathan Schaffer) states that there must be a fundamental level because otherwise "[being would be] infinitely deferred, never achieved". Trogdon suggests that the argument depends on three premises:

P1. If *A* is non-fundamental, then *A* inherits its reality from what grounds it;

P2. If *A* inherits its reality, there is some Δ which is the source of *A*'s reality;
P3. If Δ is the source of A's reality, then Δ is fundamental.

Trogdon claims that, while *P1* flows from Schaffer's conception of grounding, and *P3* flows from the same conception together with *P1*, *P2* is far more substantial, and can be supported on the basis of the following *inheritance principle*:

(IP) Necessarily, if *A* inherits ϕ then there are Δ that are a source of *A*'s ϕ-ness.

Trogdon's strategy against the inheritance argument consists in arguing that, even if this principle is correct, the property of being *real* cannot be legitimately instantiate IP. His argument seems to be that, if something inherits its reality, then that thing exists *because* of that inheritance; however, it seems that the explanatory direction should go the other way round: the fact that that entity exists helps explaining why that entity inherits some property.

Mark Jago ("From Nature to Grounding") offers an account of grounding in terms of *nature*, where an entity's nature is understood in terms of its *real definition*, and the grounding profile of that entity can be read from the logical structure of that definition. For what matters to the general aims of the collection, Jago's account supports irreflexivity, anti-symmetry, and transitivity, but fails to support well-foundedness.

John Wigglesworth ("Grounding in Mathematical Structuralism") studies the place of metaphysical dependence in structuralism in the philosophy of mathematics. Structuralists claim that *(i)* mathematical objects depend on the structures they belong to, and *(ii)* that mathematical objects depend on all other objects in the same structure. Wigglesworth shows that, if both *(i)* and *(ii)* are taken as statements about dependence, they result in counterexample to the received view; in particular, dependence relations would be non-well-founded.

Wigglesworth's paper is the most extensive study of metaphysical dependence in mathematical structuralism so far. As he points out, structuralism offers counterexamples to virtually *any* aspect of the received view; for this reason, structuralism has a prominent place in this collection. Another road-map will be useful to the interested reader. Barnes (59-60) and Thompson (118-19) mention mathematical structuralism as an example of symmetric and holistic metaphysical explanations. Litland (143) claims that structuralism provides the best motivation in favour of bicollective ground. Finally, Morganti (see below) suggests an analogy with structuralism in the philosophy of science. This collection will therefore be particular welcome by those interested in the connection between grounding and structuralism (and in the philosophy of mathematics in general).

Tuomas Tahko ("Fundamentality and Ontological Minimality") takes on a different conception of fundamentality, according to which fundamentalia are, loosely speaking, the *minimal 'reality-makers'*. This conception is nicely captured by what Tahko calls the principle of *ontological minimality*, according to which the fundamental level of reality consists of ontological minimal entities. Tahko claims, quite surprisingly, that this Ontological Minimality Principle is compatible with some forms of infinitism.

Matteo Morganti ("The Structure of Physical Reality") explores infinitism and coherentism in the light of the philosophy of science. As for infinitism, Morganti defends an alternative reply to the reality inheritance argument, according to which dependence relation do not 'transmit' existence, but existence *emerges* rather from the chain of ground as a whole. As for coherentism, he suggests that

dependence relations can be taken to be *quasi-transitive*, that is, so as to permit circles of grounding without there being any self-grounded entity.

3. The Contingency and the Consistency Theses

Nathan Wildman ("On Shaky Ground?") explores what he calls the *contingent fundamentality thesis*, according to which being fundamental is not a necessary property that *actual* fundamentalia have in every world in which they exist. Wildman illustrates how this claim can be married either to *contingentism*—the claim that whatever is fundamental is contingent—or to *necessitarianism*—the claim that fundamental entities are necessary existent, and he points out that necessitarianism faces the problem of accounting for variations among possible worlds, since everything fundamental is, according to necessitarianism, necessary existent (other relevant remarks on grounding and modality are found in Barnes' essay, 52, and in Bliss' essay, 83-84).

Filippo Casati ("Heidegger's *Grund*") formulates two *para-foundationalist* accounts. According to the first account, there is a foundational element with inconsistent properties; according to the second account, the structure itself have inconsistent properties (in particular, both E and $\neg E$ hold). He then employs these accounts to model Heidegger's notion of ground. (Para-foundationalism is not included in Bliss and Priest's taxonomy; in general, para-foundationalism is a wide uncharted land, and we would probably have liked to read more about this as well.)

The general upshots of this collections seem to be two: *(i)* grounding is best conceived as *non-irreflexive*, *non-asymmetric* and *non-transitive* rather than anti-symmetric, anti-reflexive and transitive, and *(ii)* infinitism and coherentism are lively options. At the same time, foundationalism appears to be a far more loaded position that the general agreement suggests it to be. Reading this volume is intellectually profitable and highly satisfying; so we recommend it to anyone interested in this debate.[1]

University School for Advanced Studies IUSS Pavia LUCA ZANETTI

[1] I am grateful to Ricki Bliss for sharing her work and for reading the first draft of this review.

Advisory Board

SIFA former Presidents

Eugenio Lecaldano (Roma Uno University), Paolo Parrini (University of Firenze), Diego Marconi (University of Torino), Rosaria Egidi (Roma Tre University), Eva Picardi (University of Bologna), Carlo Penco (University of Genova), Michele Di Francesco (IUSS), Andrea Bottani (University of Bergamo), Pierdaniele Giaretta (University of Padova), Mario De Caro (Roma Tre University), Simone Gozzano (University of L'Aquila), Carla Bagnoli (University of Modena and Reggio Emilia), Elisabetta Galeotti (University of Piemonte Orientale)

SIFA charter members

Luigi Ferrajoli (Roma Tre University), Paolo Leonardi (University of Bologna), Marco Santambrogio (University of Parma), Vittorio Villa (University of Palermo), Gaetano Carcaterra (Roma Uno University)

Robert Audi (University of Notre Dame), Michael Beaney (University of York), Akeel Bilgrami (Columbia University), Manuel Garcia Carpintero (University of Barcelona), José Diez (University of Barcelona), Pascal Engel (EHESS Paris and University of Geneva), Susan Feagin (Temple University), Pieranna Garavaso (University of Minnesota, Morris), Christopher Hill (Brown University), Carl Hoefer (University of Barcelona), Paul Horwich (New York University), Christopher Hughes (King's College London), Pierre Jacob (Institut Jean Nicod), Kevin Mulligan (University of Genève), Gabriella Pigozzi (Université Paris-Dauphine), Stefano Predelli (University of Nottingham), François Recanati (Institut Jean Nicod), Connie Rosati (University of Arizona), Sarah Sawyer (University of Sussex), Frederick Schauer (University of Virginia), Mark Textor (King's College London), Achille Varzi (Columbia University), Wojciech Żełaniec (University of Gdańsk)