

On the assumption of self-reflective subjectivity

History of the Human Sciences
2022, Vol. 35(2) 167–193
© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions
DOI: 10.1177/09526951211032895
journals.sagepub.com/home/hhs



Christoforos Bouzanis 
Swansea University, UK

Abstract

Contemporary social theory has consistently emphasized habitual action, rule-following, and role-performing as key aspects of social life, yet the challenge remains of combining these aspects with the omnipresent phenomenon of self-reflective conduct. This article attempts to tackle this challenge by proposing useful distinctions that can facilitate further interdisciplinary research on self-reflection. To this end, I argue that we need a more sophisticated set of distinctions and categories in our understanding of habitual action. The analysis casts light on the idea that our contemporary social theories of self-reflection are not consistent with everyday notions of agential knowledgeability and accountability, and this conclusion indicates the need to reconceptualize discourse and subjectivity in non-eliminative terms. Ultimately, the assumption of self-reflective subjectivity turns out to be a theoretical necessity for the conceptualization of discursive participation and democratic choice.

Keywords

human agency, intentionality, self-reflection, social theory, subjectivity

In this article, I *reflect* on a human property that is omnipresent – albeit on various levels, in various forms, and at various rates of intensity – in all human societies, from the so-called primitive or traditional to contemporary late-capitalist/postmodern forms. It is the prerequisite for personal decision-making in view of the existence of multiple, ambivalent, or even contradictory roles and rules in a certain situation and context; that is, a necessary element in processes of prioritization, personalization, and modification of roles and rules. It is the prerequisite for almost every kind of negotiation and act of (personal or group) resistance. It is the deeper meaning implied in our efforts to

Corresponding author:

Christoforos Bouzanis, Swansea University, School of Management, Bay Campus, Fabian Way, Swansea, Wales, SA1 8EN, UK.

Email: c.bouzanis@swansea.ac.uk

contemplate cultural differences, differences in legal systems, and conflicting social policies. Its existence as a possible everyday phenomenon of human conduct is also what we implicitly expect when we advise our students to critically analyse, compare, and present the various views found in the literature. It is what we expect to be present, or even self-evident, in most of our everyday decisions in contemporary legal systems (see Habermas, 2007), in the sense not only that we are knowledgeable of the consequences of our actions, but also that we have the capacity to critically compare and evaluate the different possibilities of action: the ideas of ‘accountability’ and ‘responsibility’ of action are premised on knowledgeability and critical evaluation of the possibilities of action in most of the contextual and situational settings we find ourselves in. It is also what is supposed to express the commonsensical view of democratic choice: that those who have the right to vote are expected to choose – by critically weighing – candidates, parties, and ideologies. It is what we are supposed to be doing all the time in academia: critically reflecting on models, theories, methodologies, data sets, arguments, and counterarguments.

And yet, while *self-reflection* is theoretically and praxeologically *essential* for agential conduct, it remains *undertheorized* in contemporary social theory. Of course, there have been a few influential attempts to focus on this phenomenon (see e.g. Archer, 2003; Wiley, 1994), and admittedly the references to ‘self-reflection’ or ‘agential reflexivity’ in contemporary literature are, one could claim, relatively frequent, but most of the references to human reflectivity are sporadic and/or only a minor part of broader projects. Therefore, one could conclude that even though many key authors make clear reference to self-reflection, there remains a relative reluctance to offer an in-depth analysis of it. For simplicity’s sake, I will treat these two terms (*self-reflection* and *agential reflexivity*) as synonymous and mutually interchangeable, though *self-reflection* is more in harmony with its use by classical social theorists like George Herbert Mead, as well as some key authors in contemporary social theory (see Emirbayer and Mische, 1998; Porpora and Shumar, 2010); this use of *self-reflection* also coincides with its use in different key philosophical debates (see e.g. Habermas, 1996; Joas, 1996; Sosa, 2009) and also in certain key analyses in social psychology (see e.g. Benbassat, 2020; Martin and Gillespie, 2010).

Yet some of the key authors who have dared to touch upon this difficult-to-grasp phenomenon, and have made a contribution to our understanding of it, seem to have confused the discussion not only terminologically but also by blurring key distinctions that might facilitate a more empirical investigation and a more adequate understanding of the phenomenon. Norbert Wiley, for example, who offers us an impressive overall analysis of self-reflection, discusses heterogeneous ideas by placing them all together under the same homogenizing umbrella term, ‘Reflexivity’ (Wiley, 1994). That is why, in order to discuss self-reflection (agential reflexivity), we need to draw upon key analytic distinctions. First of all, we need to distinguish between self-reflection (agential reflexivity) and epistemic reflexivity (Bouzanis, 2017; Bouzanis and Kemp, 2020). In this article, I am focusing on the former, although I will occasionally refer to the latter when necessary.

Moreover, as I will explain in the first section, though omnipresent in the various micro-contexts of every social world – even in the so-called traditional societies – self-reflection appears at different levels, in different forms and modes of processes, and at different rates of intensity. In addition to this multidimensional character of its

omnipresence, and as I implied in the first paragraph, self-reflection constitutes an assumption, a prerequisite, an expectation, or a commonsensical and self-evident accomplishment in various (sub)fields or contexts of our contemporary late-capitalist societies: academia, legal systems, democratic decision-making, negotiations, and employment relations. Thus, this article aims, at the first level, to make these distinctions even clearer.

This article also intends to cast light on a common problematic in the analysis of self-reflection in social theory. I have mentioned that references to self-reflection in the contemporary literature accord it only a peripheral theoretical status and are rarely accompanied by any in-depth analysis. The word *peripheral* implies something important: that in many influential analyses in social theory, self-reflection is inadequately related or even excluded by the interrelated terms of the core parts of the various theoretical frameworks, or by the assumptions of the related social ontology. This article intends to turn the tables. Instead of starting from a single social theory and assessing it in terms of whether this theory offers a persuasive analysis of human reflectivity, or whether this theory (explicitly or implicitly) excludes such an analysis, the chosen mode of argumentation takes a backward direction; it starts from generic, analytic distinctions and remarks relating to the understanding of self-reflection, and then asks a key question: ‘What can these generic distinctions and remarks tell us about our best theories that are supposed to highlight self-reflective conduct as an agential capacity?’

The puzzling conclusion would be that contemporary theories of self-reflective subjectivity are not in harmony with the commonsensical, everyday assumptions of knowledgability and accountability of action. The more promising conclusion is that by starting our tracing of human conduct from the assumption of self-reflective subjectivity, we can set restrictions on theoretical speculation, as human reflectivity turns out to be both a core ontological assumption and a *fact* of our personal and shared histories.

In the first section, I will discuss the multidimensional character of self-reflection in order to identify essential distinctions and remarks. In the second section, I will briefly examine how we can reconsider non-reflective human conduct in relation to self-reflection. This analysis will help us in our broader effort to place self-reflection within social theory in the third section, where I follow an imaginary, bifurcated path through the key moments in the theoretical construction of self-reflective subjectivity – moments that result in part from the analysis of the distinctions of the first two sections – and I place heterogeneous theories of self-reflection in each of these moments, so as to assess whether any of these theories are in harmony with assumptions about the self-reflective agent.

Self-reflection: Preliminary distinctions

There is some confusion about how we can study self-reflective moments, let alone how we can achieve a deeper understanding of self-reflection, due to the absence of key analytic distinctions that would potentially help social research to avoid unnecessary generalizations or overlaps. On this point, we need to draw on the four dimensions of how to understand human reflectivity: levels, forms, modes of processes, and rates of intensity.

Levels of self-reflection

We need to distinguish among fallible reflection on: (a) inner states; (b) others' mental states and meaningful actions and statements; (c) micro-contexts of situational interaction; and (d) macro-domains of sociocultural interaction. The first two levels are the ones that most of us understand from Mead's and Cooley's early approaches to more contemporary formulations: that is, that we have the capacity to objectify and communicate certain (though not all) inner mental states, and also that we are able to (fallibly) attach interpretations to the actions and sayings of others that are directed towards us.

But we can also (fallibly) objectify actions and sayings that are not clearly directed towards us (see Gillespie, 2006: 230–69); we can meaningfully objectify actions that are clearly not directed towards ourselves; we can (fallibly) objectify micro-contexts we find ourselves in, and macro contexts, such as social structures and global or regional cultures and institutions, that somehow relate to our conduct. At the heart of the discussion about the modes and levels of human reflection, we face the problem of how to distinguish between our representations of objects in the subworld(s) we participate in, and self-reflection that denotes a (frequently critical) objectification of one's own intentions, beliefs, values, images and imageries (complexes of images), and other mental states (on this, see Von Wright, 1992).

So, at this point, it will help our analysis to *reflect* on the idea of objectification: this concept is broader than reflectivity, since we are able to formulate an image of an object of perception without generating a kind of cognitive dissociation (Rafieian and Davis, 2016) or distance (Gillespie, 2006) from the image as such. It is this hypostatized cognitive distancing between the 'I' and an image or imagery that constitutes the genre of reflectivity, rather than the broader and more generic notion of objectification that can equally entail imaging without distantiating.

Forms of self-reflection

Here, the list of forms can be open-ended, depending on which one thinks are the most basic or generic forms of expression of reflectivity: forms of description, affirmation, critique, and comparison; followed by deconstruction, constructive theory production, deduction, abduction, and retroduction. These are just a few examples of forms of understanding of our distantiating from images of events, objects, and others, and from imageries of self/others and of micro and macro contexts. The literature on these forms is exhaustive, and the reader can relate each level of reflection to different forms.

Yet it is fruitful to focus on the notion of deconstruction, since its frequent use has generated a deep confusion in social theory, related to the common error of conflating the social scientist's self-reflection with epistemic reflexivity. It is no exaggeration to claim that the history of the notion of reflexivity in epistemological circles is a series of confluences of philosophical and scientific self-reflection with epistemic reflexivity (see Bouzaris, 2017).¹ The former designates all those efforts to objectify (a) one's own beliefs, values, assumptions, ideas, schemes, methods, models, and other kinds of epistemic premises; (b) laboratory processes and institutionalized micro-contexts that one finds in academia; and (c) shared ideational backgrounds and norms, macro-

institutional settings, and structural settings of the academic world(s). The latter designates the theoretical property of auto-reference that pertains to grand social theories and social ontologies: 'the application of socio-theoretical systems to their own academic microcosm, such that the general characterizations of social relations put forward within a theoretical system have an (explicit or implicit) import for the analysis of the conditions of theory-production themselves' (Bouzanis and Kemp, 2020: 66).

The form of conflation between these two analytically distinct, though intrinsically intertwined, notions depends on the epistemological standpoint that is related to the (social) theory of knowledge utilized by each specific theorist of reflexivity. For example, from a Bourdieusian, objectivist standpoint, the conflation takes the form of objectifying 'the subject of objectification in its objectively structured academic microcosm' (Bouzanis and Kemp, 2020: 73); in ethnomethodological, constructionist, and postmodern approaches that underline the textuality of accounts, the conflation has taken either the form of the celebrated instability of meaning that is emergent in the auto-referential application of forms of epistemic relativism to its own premises (Ashmore, 1989; Woolgar, 1988; Woolgar and Ashmore, 1988) or the form of the rhetorical confession of one's own standpoint by uncovering epistemic values or ontological commitments (Pels, 2000; Pollner, 1991). In this latter tradition of deconstructionism, the self-reflective uncovering of one's own commitments is conflated with the auto-reference of the idea of textuality of accounts to its own epistemic premises. But again, we first need to analytically distinguish between the two in order to show their interrelations (see Bouzanis, 2017).

Modes of processes of self-reflection

In this section, I cast light on the processual character of self-reflection by pointing to the paradigm transition from introspection to self-dialogue, in which mental states are scrutinized in and through inner dialogues.

Allow me, at this point, and before moving forwards, to clearly underline the poverty of our metaphors of self-reflection (agential reflexivity) in our early efforts to understand the 'mechanisms' and processes of this enigmatic phenomenon. Introspection and self-dialogue, of course, are not mere metaphors; they are *real* expressions of the broader phenomenon of self-reflection – for, after all, our reflective imagination is expressed in visualizations and scientific modelling, in which iconic imaging is the key feature (see Bouzanis, 2016; Harré, 1975), and day-dreaming is usually expressed in inner dialogues. But we are still lacking a platform theory (combining inputs from the various human sciences) of how to incorporate introspection and self-dialogue into a generic ontology of self-reflection.

All in all, Gillespie (2006: 267) is timely in claiming that we do not know a lot about the exact mechanisms of distanciation in self-reflection. Nonetheless, self-dialogue is replacing the idea of introspection in social theory, as influenced – among other key tendencies, such as the linguistic turn in the human sciences – by the resurrection of self-talk in Margaret Archer's critical realist theory, which has granted self-reflection a key role in the definition of agency. Archer offers an analysis of agential reflexivity as the mediator in the causal interplay between the restricting and enabling powers of social structures

and the human powers of imaginative agency to reflect on inner states and outer settings – taking the form of the ‘Internal Conversation’, which, as an approach, respects ‘the interiority, subjectivity and causal efficacy of the life of the mind’ (Archer, 2003: 93).

For Archer, inner dialogue is a preferable notion to introspection or self-perception because, unlike the latter, it acknowledges that we can talk and listen to ourselves simultaneously; here, we have two functions working at the same time, whereas in the case of introspection, we make use of only one of our senses, that is, perception. For Archer, we internally discuss with ourselves in order to reach (not infallibly) self-knowledge and achieve self-modification. Of course, ‘like all conversations, it is one that takes place over time’ (Archer, 2003: 97), and it is one that involves turn-taking. Archer describes this inner dialogue as an alternation between subject and object through which the monadic self critically responds to the objectified echo of her previous statement: ‘The two will go on alternating until solidarity is reached or the issue is postponed or abandoned’ (ibid.: 100).

The application of this process is as follows. Through internal dialogue, we endorse, review, choose, or deny our commitments, and we also learn about them. It is also through this internal deliberation, through the (fallible) consideration of our enablements and constraints, that we can formulate projects, which are crystallizations of our ultimate concerns and based on which we (fallibly) expect to achieve, through an objective structural context, what we really want. Thus, the interplay between two kinds of causal power (structural and agential) is feasible through the Internal Conversation, and the activation of structural powers is due to the formation of these projects. If the subject changes her ultimate concerns (which constitute her personal identity), that is, changes herself, personal morphogenesis will be the case, and if she continues to endorse those concerns, we will have personal morphostasis. Whether we have morphogenesis or morphostasis in the social world depends on the (unpredictable, irreducible, and non-determined) interaction among agents who have formulated various projects.

However, Archer’s account of how structure and agency integrate into one story is not a plausible one. Agential reflexivity is supposed to be the mediator of sociocultural constraints and enablements. But Archer sometimes refers to ‘the interplay between sociocultural properties and the exercise of agential reflexivity’ (Archer, 2003: 130). Human reflectivity as a key dimension of human subjectivity frequently appears to be both the causal mediator and a causal power in a causal interplay. Why is this so problematic? The mediatory process is a causal power in an interplay that is simultaneously mediated by it. The problem here is that the interplay is assumed to involve two distinct powers, at the same time that one of these necessarily mediates the other in their mutual interplay; questions inevitably arise about the nature of mediation, as well as about whether it is agency or structural powers that have the upper hand in this story of interplay (on this, see Kemp, 2012).

Moreover, according to Archer (2003), we formulate our projects in order to achieve what we ultimately want. The Internal Conversation is the process that leads us to these projects, while our concerns are the main kind of input in this process. But when we apply the Internal Conversation to the decision about our ultimate concerns, Archer is not clear about the ontological status of the inputs that enter into this process. The only clue Archer gives is what she borrows from William James, namely, ‘thought tendencies’: ‘The

subject, he insisted, has a “premonitory perspective” on what she is about to think. This means, I take it, that she has both a blurred notion of an uncrystallised thought, and some inchoate awareness of what its thrust should be’ (ibid.: 98). And, at this point, we may ask what kind of cognitive property it is that enables pre-linguistic understanding. Alternatively, do ‘premonitory perspectives’ emanate in some way from shared cultural perspectives?

Additionally, Archer continues, after we have selected the appropriate words, this articulation is the first utterance that will become an object, which the subject can question or accept. She can then reformulate or abandon this effort; for ‘perhaps what she hears she immediately realises cannot be correct, because listening to the utterance also brings counterfactuals to mind’ (Archer, 2003: 98). But are these counterfactuals crystallizations of other premonitory perspectives? The answer could be yes since, as Archer explains, in this lengthy process of internal discussion, ‘we have to inspect, interrogate and evaluate the object in relation to a premonitory thought tendency towards it, and then formulate a new utterance in response to it’ (ibid.: 106).

Archer (2003) does not connect these ‘thought tendencies’ with the ‘Me’ (past self), which the ‘I’ (the present self, who is the only one responsible for the questioning of concerns and for the conversation) always takes into account; that is, she is not clear about whether these tendencies come from the past, the past that, in its turn, conditions the present self. In such a case, she might invoke memory, but not as an interlocutor, which would be impermissible for her. Indeed, this was her critique of the supporters of retrospection and of William James, for ‘memories cannot be interlocutors’ (ibid.: 95).

And we cannot find answers in Archer’s *Being Human* (Archer, 2000), where the inner discussion occurs between the ‘I’ (the present spontaneous self) and the ‘You’ (the future prospecting self), both of which are constantly informed by a (non-Meadian) ‘Me’ formed of ‘all the former “I”s who have moved down the time-line of future, present to past’ (ibid.: 229) and thus constituting a memory store of past emotional commentaries, providing data for the two interlocutors (the ‘I’ and the ‘You’). Yet it is now more than obvious that this account is not in harmony with the one in Archer’s *Structure, Agency and the Internal Conversation* (Archer, 2003), where such a division of the self is unacceptable – for, although it is true that the past and the future exert influences and constraints on the ‘I’ while replying to himself, ‘these processes should not be reified by endowing the past “Me” or the “You”-to-be with the power of speech or of hearing’ (ibid.: 111).

The main argument here is that Archer’s Internal Conversation seems to be an analytically *useful* process. However, Archer is unclear about the process of this internal dialogue, about the interlocutors, and about the origins of the cognitive elements that appear as inputs in this process.

This, I think, is a result of the fact that Archer is ambiguous on the cultural origins of people’s evaluations (see also Farrugia and Woodman, 2015). Now, though her ontological distinction between agency and culture is a useful tenet, Archer (1996) understands culture as consisting of objective items, texts, and logical relations. Yet to distinguish subjective meaning from shared ideational backgrounds does not *necessarily* resurrect Popper’s (1972) third world; we can instead discuss the possibility of distinguishing subjective meaning from (theoretically reconstructed) intersubjective ideational

backgrounds and explain how the possibility of reflective subjects is a prerequisite for understanding a participation in discourse – in which ideational backgrounds are modified by self-reflecting agents. This is the final goal of this article, discussed in my concluding remarks.

Another important view of the inner dialogue as an expression of self-reflection has been provided by Wiley (1994, 2006a, 2006b, 2010), who attempts to provide an approach that unifies Mead's 'I-me' with Peirce's 'I-you' version of inner speech. The 'I-you-me' triadic notion of reflexivity that he proposes constitutes a combinatory model in which the 'I', the 'You', and the 'Me' should be considered more like complex relations than substances.

One can envision both the past and the future, along with the present, simultaneously. This omni-scopic vision allows one to go back and forth from past to future and also to see them together. If one can simultaneously see one's settled habit system ('Me') and one's options for some new, non-habitual action ('You'), one can more easily integrate the two practical resources, structure and agency (Wiley, 2010: 18).

Here, Wiley's 'Me' as habitual system is not similar to Archer's 'Me', since Wiley attributes to the 'Me' component five partly heterogeneous contents. These are the 'generalized other', 'habits', 'memory', 'interface with body', and 'self-concept', and 'there is no reason why all three cannot take turns speaking, and for the dialogue to be between the "me" and "you", as well as between the "me" and "I" (or the "you" and "I")' (Wiley, 2010: 20). For this triadic relationship to occur, the self must stand simultaneously in the past ('Me'), present ('I'), and future ('You').

According to Wiley, since *existentially* we can stand only in the present, we can achieve this multilevel interplay imaginatively (Wiley, 2010: 20). He underlines the extra-linguistic dimensions of inner speech, as well as the role of imagery in our dialogical thinking:

If we remember that inner speech is imagistic as well as linguistic, a principle of differentiation is evident. This is in the elaborate batch of imagist materials we can use for parts of speech. Since these are only loosely tied to the meanings they might represent, they can vary quite a bit. I would suggest that the words of inner speech tend to be standardized and small in number, but the imagery is less standardized and larger in number. (Wiley, 2006a: 332)

Wiley does not distinguish between images and imageries, but what is important here is that he is clear that 'the imagery, primarily visual but also tactile in this case, is dominating in the inner language. These images function as various parts of speech, but they also function as signs of external objects' (Wiley, 2006a: 331).

Reflecting on the differences between these two approaches to inner dialogue, Wiley (2010) enhances Archer's (2000) initial remarks on the 'I'-'Me'-'You' interrelations that are absent in Archer (2003) and allow a more advanced role for imagination in our creative responses to our imaginary interlocutors, and also our own previous positions in dialogic relations with others. From my perspective, what is missing in both approaches is the recognition of the reflective subject as both the participant who internalizes elements

from ideational backgrounds and the agent who discursively modifies not only structural settings but also world-imageries and normative orientations.

Self-reflection appears at different rates of intensity

In the introduction, I suggested that self-reflection is assumed – more or less, depending on situational logics and institutionalized expectations – to be a self-evident or essential assumption in many of the everyday fields and contexts that we participate in in our everyday conduct. And yet, a number of key contemporary authors frequently locate the emergence of self-reflection in the tendencies of ‘reflexive modernization’ (Beck, Giddens, and Lash, 1994), an era in which individuals have no alternative but to contemplate choices. And this is frequently counterposed to the so-called ‘traditional societies’, where routine action, role-performance, and rule-following was the *rule*, instead of critical reflection on habitual dispositions, roles, and rules. Yet Archer is *partly* right to claim that agential reflexivity, even in traditional societies,

was not entirely absent and could not be because it is socially indispensable in three ways: a reflexive ‘sense of self’ is necessary for the correct appropriation of rights and duties by those to whom they are ascribed, the self-monitoring of performance is necessarily a reflexive task, and reflexivity is crucial for bridging the gap between formal expectations and actual eventualities in the open social system. Nevertheless, what reflexivity does not and cannot do in traditional societies is to enable its members to re-envision either the self or the social because they lack the ideational and organizational resources for doing so. (Archer, 2012: 60)

I say ‘partly right’ because, while we need to recognize that only certain discourses promote and allow the critical scrutiny of self-images and discursive positionings (Harré and Gillett, 1994: 127; see also Smith, 2002), critical reflective conduct – re-envisioning either the self or the social – was always there even in the most cohesive and stable sociocultural systems of meanings – for otherwise, sociocultural (intentional) change would not be possible. But this is the subject matter of the final section. First, let us briefly *reflect* on certain tendencies in social theory relating to how we can combine reflective and non-reflective conduct.²

Reflective and non-reflective conduct

Two opening remarks should be made at this point. First, many of our best social theories exclude or tacitly neglect or downplay the possibility of self-reflection, by overemphasizing habitual action (Bourdieu, for example), or rule-following (the Wittgensteinian tradition, for example), or role-performing (Goffman, for example). Second, key social theories that clearly incorporate self-reflection as a possibility in the repertoire of the social actor have usually tried to relate self-reflection to other modes of action. This is a reasonable tendency, as long as we are not trying to reconcile the irreconcilable.

In this sense, many authors (Crossley, 2003; Decoteau, 2016; Sweetman, 2003; Taylor, 1995) have tried to reconcile Bourdieu’s notion of habitus with self-reflection,

an effort that has resulted either in a lack of ontological depth to the processes of self-reflection or in the assumption of paradoxical, self-contradictory, hybrid forms like the idea of 'reflexive habitus' (on this, see Archer, 2007, 2010a; Bouzanis and Kemp, 2020). Yet, due to the importance of Archer's approach to self-reflection and Bourdieu's analysis of epistemic reflexivity, other authors, like Elder-Vass (2007, 2010), have striven to offer a combinatory ontological approach that encapsulates the best of both worlds (see also Mutch, 2004).

While Elder-Vass (2007, 2010) partly realizes the difficulties of such an effort to reconcile Bourdieu's notion of the habitus and Archer's notion of reflexivity, he proposes a combinatory account in which agents make conscious decisions that are only indirect and partial causes of their behaviour, in that '(a) they occur a variable length of time before the action concerned; and (b) they are always incomplete regarding the details of the action to be taken ... thus any single human action may represent the (full or partial) realization of a series of nested decisions of various sizes or scope' (Elder-Vass, 2007: 339).

In this way, previous decisions become inputs in the process of consciously making current decisions, all of them different in size, scope, and effect. Old conscious decisions are stored as dispositions, and thus, in the end, action implementation is determined partly by those dispositions – including dispositions stored as such by 'last minute' decisions. This is a formulation in which 'our conscious decision making and our nonconscious behavior determination appear as complementary and mutually necessary moments in the causation of our actions' (Elder-Vass, 2007: 340).

Putting aside the fact that Archer (2010a) thinks that Elder-Vass misinterprets Bourdieu and simplifies her own account of agential reflexivity, and leaving aside the interesting question of whether Bourdieu really allowed for levels of self-reflection (see Archer, 2010a, 2010b; Bouzanis and Kemp, 2020), it is important to state at this point that the expression *conscious decision* generates more issues than is helpful, because self-reflection is assumed to be the socio-theoretical *alternative* to outdated philosophical versions of cogito(s) and, although sophisticated, Elder-Vass' approach tacitly conflates routine action and habitual action. Non-reflective action is not exhausted in Bourdieu's habitus, let alone in Elder-Vass' notion of dispositions, and social reproduction is not necessarily the result of non-reflective action. Let us explain this.

Before we reconcile seemingly heterogeneous forms of action by offering a general theoretical framework for subjectivity and selfhood, we need to draw on useful distinctions. First, we need to distinguish between (a) routine action; (b) habitual action; and (c) embodied dispositions. The first of these designates the repeated patterns of institutionalized and coordinated interactions that are practically shaped by the hierarchies of roles and the matrices of rules in various fields of action. We cannot claim that our routines *clearly* belong to non-reflective action: we can follow an organizational rule while simultaneously reflecting on our view that it is nonsensical, useless, catastrophic, incomplete, or promising. After all, clarification and personification of, and selection among, various social roles are conducted in and through self-reflection. Habitual action refers to the non-reflected *personal habits* of everyday life (e.g. I drink orange juice before I get dressed every morning), which are subject to minor or major change after reflection and the intention to modify them. Elder-Vass' model pertains to this second modality. We also embody non-reflected dispositions (the third modality listed above), which, while

culturally and institutionally internalized, can vary in their level of persistence (I cannot easily change my accent, but I can relatively easily change my cultural tendency to offer my hand when introducing myself if this is suggested by epidemiologists). Though there is much overlap among these three modalities, we are nonetheless able to offer ideal-typical formulations of them.

We also need to state that routine action and non-reflective conduct should not be considered guiding factors towards sociocultural reproduction. This is a deep and disorienting misunderstanding that has its main origins in Bourdieu's conflation of routines, habits, and embodied dispositions in the concept of the habitus. Whether ideational backgrounds, institutional settings, or structural settings are *intentionally* and/or *reflectively* modified or sustained, or whether they are unintentionally modified or sustained, this cannot be an outcome that is dependent on our routines, habits, or embodied dispositions (at least not until it is adequately explained). I will *reflect* on the latter further in the final section.

At present, many social scientists from various backgrounds can agree that it is legitimate to try to offer a unificatory, reconciliatory approach that combines self-reflection with our routines and non-reflective states (habits and dispositions), as well as with other concepts that denote key themes in sociology of the self, such as emotions (see e.g. Burkitt, 2012; Holmes, 2010). Elder-Vass' interesting model can be further utilized and extended in this direction, but again, we need to be careful in how we define the elements that are to be integrated. Finally, every effort to combine these elements with self-reflection needs to transcend any individualist echoes in recent conceptualizations of self-reflection (as in the Archerian acultural model) and show how ideational and cultural backgrounds are related in a dialectical way to processes of self-reflection.

In this respect, one of the goals of this article is to set the basic lines of how the assumption of self-reflective subjectivity is a prerequisite (a theoretical necessity) for discursive participation.

Our best social theories and the (im)possibility of political praxis

I have discussed four dimensions of (the sociological analysis of) self-reflection (agential reflexivity) and suggested that we need to introduce a diversification of our conceptual apparatus for what we understand as routine and habitual action. In highlighting the fact that many of our best social theories exclude, tacitly neglect, or downplay the possibility of a subject who can reflect on inner states, outer settings, and shared backgrounds, I have also examined certain dominant theories of self-reflection, noting that they have excluded or downplayed culture and ideational backgrounds in their combinatory effort to offer a unificatory model of how self-reflection relates to habit. I have indicated that we need something more than this. We need to introduce cultural and ideational backgrounds into our integrative, interdisciplinary analyses.

Yet one could respond at this point that even if key authors in the critical realist tradition – who also draw on American pragmatist views of the self – have focused on combinatory efforts that downplay the role of shared cultural and intersubjectively shared ideational backgrounds in the analysis of the construction of selfhood, constructionist and post-structuralist approaches to the self have overemphasized the discursive and

cultural shaping of subjectivity, to the extent that the expression *self-reflective subjectivity* seems to be a contradiction in terms, or at best a paradoxical creature that is assumed to combine two mutually contradictory elements. But is this an inescapable dilemma in social theory? That is to say, must we either exclude or downplay the possibility of self-reflection in our analysis, or rather downplay the idea of discursive elaboration of self and identity? I think not, as I explain below.

First of all, contemporary pragmatist approaches have emphasized the situatedness of habituated action (Strand, 2019) and interaction, and the parallel general absence of reflectivity and formation of world-images as the background of reflection. For example, Strand and Lizardo (2015: 46), contra interpretivism and speculative philosophy, argue against the idea of sharing culturally shaped world-images as guiding principles of belief formation and action. Here, embodiment, habituation, and non-representational conditioning of situated practices prevail in principle over the *reflection* on shared norms or world-images.³ For Strand and Lizardo, there is no reflection prior to action (ibid.: 48, 53), and human reflectivity emerges – in a Bourdieusian fashion – only in ‘exceptional circumstances’ (ibid.: 53), such as unusual moments of mismatch situations (Strand and Lizardo, 2017). Yet, as is highlighted by Bouzanis and Kemp (2020), Lizardo’s (2004, 2021) position is that habit is the dominant mode of action in social life, and the *flexible* habitus is the main cognitive resource that we have in our efforts to cope with modes of mismatch or hysteresis, that is to say, with situations where there is no correspondence between dispositions and structural conditions (Lizardo, 2004: 391–2; 2021: 3). Nevertheless, in these cases, habit-based explanations are the only game in town (Lizardo, 2021: 7–8).⁴ As a consequence, habit-based explanations theoretically exclude or replace representationalist and ‘mentalist’ (ibid.: 10) ascriptions of intentionality (ibid.: 14), the ‘reflexive contemplation of ends’ (ibid.: 17), and ‘reflexive intentional action’ (ibid.: 16).

Additionally, the dominant tendency in the social constructionist account of the self is the rejection of the assumption of interior, fundamental psychological processes (Gergen, 1985: 271) in favour of the idea of discursive action as relationally embedded (Gergen, 2011: 113). Here, the concept of the self as performance (Gergen, 2009) implies that there are no inner substances or processes but that the self is following scripted sets of interdependent, contextual interactions that are coordinated by language. Thus, action is meaningful only as long as it addresses certain audiences using different languages in various contexts.

At this point, we can recall that the notion of subjectivity is closely related to ideas of the discursive production of identity and of the linguistic/symbolic character of the performance and presentation of the self – especially in *certain* works of Michel Foucault and Judith Butler, where it is relatively clear that the idea of the discursive production of the self excludes any notion of the internalization of shared ideas, ideals, and norms. This in turn means a definite negation of the possibility of self-reflection, for subjective indexical expression remains meaningful only at the surface of discursive practices. There is no ontological depth of subjectivity, and any effort to discuss a dialectic between subjectivity and intersubjective shared ideational backgrounds is dismissed as another expression of the ideology of the individual.

However, no matter what Foucault really meant by the idea of discourse (see Caldwell, 2007), we still need to disentangle the idea of discourse as symbolically mediated interaction from certain other dimensions that permeate the social realm, such as material conditions or non-discursive practices (for an example of this, see Alvesson and Kärreman, 2011). It is also important to note that while the idea of subjectivity is not necessarily linked to the idea of discourse – for a number of key authors in social theory have utilized the former without the latter – it is reasonable to claim that the idea of discourse is inextricably linked to the idea of the sociocultural production of subjectivity. It is at this point that we need to analytically disentangle these two ideas in order to show their interrelations. If we cannot, it would appear that the idea of ‘self-reflective subjectivity’ is indeed a contradiction in terms, for the subject would not be able to distance herself from routines, habits, scripts, and/or performances.

There are a number of interesting voices all suggesting that we need to analytically distinguish the notions of subjectivity from discourse when trying to produce a more complete story of their theoretical relation. González Rey (2019a, 2019b), for example, has recently proposed an idea of subjectivity that emphasizes the roles of emotion and imagination in the unpredictable creation of different microcosms (González Rey, 2019a). Subjectivity here, as a quasi-entity that is non-reducible to language and discourse (González Rey, 2019a: 184; 2019b: 226–7), is both fluid and relational, in the sense that emotions attain a symbolic character in becoming subjective senses that are intertwined and integrate with each other, formulating dynamic chains from which subjective configurations emerge – configurations that are not determined by external motives, nor are they the internalized impact of linguistic expressions.

Unlike in the constructionist tendency to present the self as the production of inner or outer dialogues, here self-dialogue is possible because of this emergence of subjectivity (González Rey, 2019b: 230). For González Rey (2019a, 2019b), the subject is not acultural – there is interchange with social figurations – and he also states that

individuals and social institutions are not external to each other. Both of these systems have their own subjective configurations but, at the same time, constitute each other, not as a result of the external influence of one upon the other, but through the subjective senses generated by each of these configurations during a specific experience. (González Rey, 2019a: 183)

Yet this version of ‘weak’ mutual constitution between two different kinds of subjectivity (i.e. individual and social) is not very clear or helpful in our effort to formulate an ontological description of how subjectivities and discourses relate to other elements of social reality. However, the distinction between individuals’ subjectivities and institutional roles and rules is essential to understanding the idea of a self-reflective subjectivity that can distance herself (at a theoretical level, in a normative way, and in practice) from institutional demands and material conditions. We need a notion of subjectivity that is not conceived of as a shadowy expression of local practices or a cultural dupe that merely reproduces internalized representations and normative directions, but that rather highlights a deeper understanding of how the subjective internalization of intersubjective shared ideational backgrounds is one part of a dialectic with the possibility of self-reflection – and also recognizes that participation in discourse has self-reflection as a

prerequisite idea. For this, in the future, we should move beyond the old distinctions of pronouns ('I'–'Me'–'You') and thus make an interdisciplinary effort to discuss subjectivity from various angles.⁵

In what follows, I will be utilizing the above-mentioned distinctions and themes in order to ask what the distinctions we have established so far can tell us about our best social theories of self-reflection. I need to be clear that what follows is not meant to be another critique of the idea of the 'oversocialized conception of man' (see Wrong, 1961), an idea that frequently, in its variations (as I have briefly sketched out in the previous paragraphs), excludes the possibility of self-reflection. Instead, I focus on those key theories that have incorporated ideas of inner dialogue and self-reflection – that is, theories that more or less assume the possibility of self-questioning – and try to show how they have presented the self-objectifying subject in relation to other interrelated assumptions. This will lead us to the conclusion that dominant theories of self-reflection are not in harmony with commonsensical accounts of knowledgeability and/or accountability of action.

In order to do this, I have distinguished four key questions that are crucial to the theoretical understanding of self-reflection, and which have been more or less implied at various points of the above analysis. These questions will turn out to be dilemmas in theory construction. We will proceed to this analysis of self-reflection in order to discuss the repercussions of its various (and heterogeneous) conceptualizations for the theories (from radically different traditions) that incorporate these conceptualizations.

Q1: Does self-dialogue necessarily entail human reflectivity?

From Vygotsky's dialectics and Bakhtin's dialogics to more contemporary formulations of the dialogical self, self-dialogue is a possible *modus operandi* of self-reflection.⁶ However, the latter is not *necessarily* exhausted in the former. In fact, one could claim that self-reflection is ontologically the more basic, broader phenomenon, and is in need of further investigation. So far, so good; I think this is quite intelligible. However, the inverse question is whether formulations of the dialogical self *necessarily* imply self-reflective conduct, even in the minimalistic form of self-monitoring (as in Giddens, 1984).

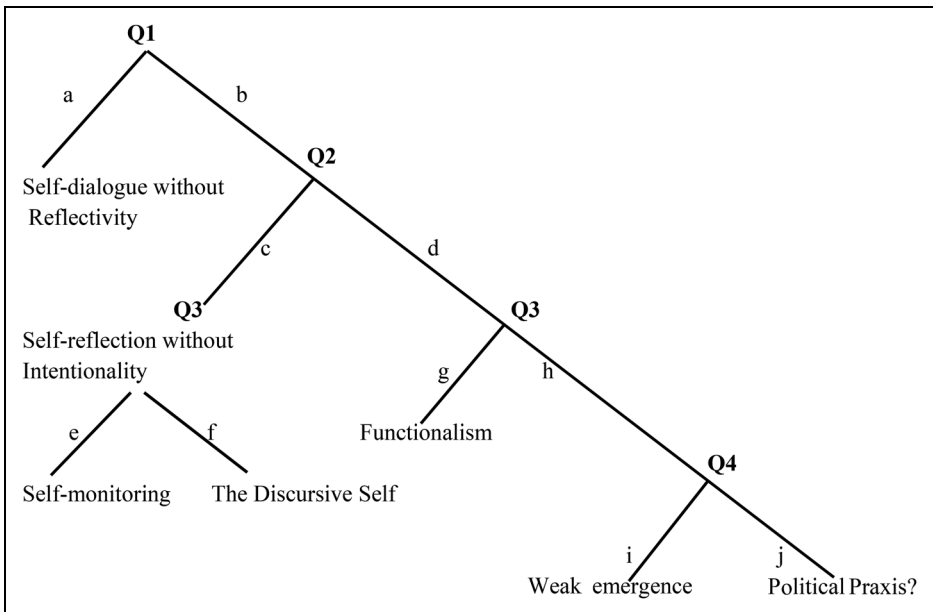
We can see that the answer to our question can be negative; there are formulations of the dialogical self that do not allow any kind of distantiation in the sense referred to above. Just to take an example, drawing on the work of Mikhail Bakhtin and William James, Hermans (2001, 2002, 2003, 2006, 2012) proposes the idea of the polyphonic decentralized self, a multidimensional and dynamic self, moving dialogically within an imaginal, collectively negotiable, and dynamic space of multiple (inter)cultural and trans-personal position-takings of the 'I'. That is to say, a self that is moving imaginatively within an ever-reconstructed field of multivocality that defies the esoteric/exoteric distinction. For Hermans (2002: 148), 'The *I* fluctuates among different and even opposed positions, and has the capacity imaginatively to endow each position with a voice so that dialogical relations between positions can be established'. Each position constitutes 'a character' or another 'I' with a respective 'Me'.

Hermans argues for the ontological complicity between society and the self: they are both characterized by relations of agreement, disagreement, tension, contrast, or dominance among different interlocutors placed at certain positions. The theory of the dialogical self as a society of minds or as a theatre of voices assumes that ‘there is no essential difference between the self as a society of mind and the “real” external society’ (Hermans, 2006: 150). Yet this mode of ‘strong’ mutual constitution demolishes the boundaries between the ‘I’ and ‘Others’ and tacitly excludes reflective conduct.

Hermans makes many references to Mead. No matter whether we take Mead as one of the first key authors who theorized the possibility of self-reflection (Emirbayer and Mische, 1998; Martin and Gillespie, 2010: 257) or as the author who provided an over-socialized version of the self (Archer, 2003: 78–90), which would make self-reflection impossible, we still need to state once again that the ‘I’–‘Me’(–‘You’) distinction is a crude formulation of the problem and that we are in need of a more interdisciplinary collaboration in taking the next step in the conceptualization of self-reflective conduct.

Q2: Is the idea of self-reflection accompanied by a notion of intentionality?

We can now trace our imaginary path in theory construction, following key theoretical decisions, in each stage of figure below. Each move leftwards signifies an (explicit or implicit) negative stance towards the respective questions; accordingly, each move



rightwards signifies an (explicit or implicit) positive stance towards the question posed each time.

Consequently, (a) leads to conceptualizations of self-dialogue that negate the possibility of reflection, and (b) leads us to the second key question, that is, whether in understating human agency we need ideas of both self-reflection and intentionality.⁷

In this sense, having followed (c), we can find theories that discuss a reflective mode of 'thought as bending back upon itself' (Tauber, 2005: 50), but any kind of intentionality is unintelligible, at least according to a contemporary understanding of it. I cannot find any better, or less extreme, example than Blum and McHugh's (1984) ethnomethodological approach to self-reflection. Here, language as the unification of convention recollects its own intelligibility in actors' speech, an idea that leads us to self-reflection as a retraction of the centrality of language in the fragmentation of speech. The non-intentional subject becomes a shadowy quasi-entity that appears in the mysterious, ironic, self-reflective moments when language as an ultramundane totality identifies itself as pervasive convention.

But let us now look at a more influential approach that I think combines both self-reflection and intentionality, though this is debatable, as are most of the categorizations in this third section. In Harré's work (Davies and Harré, 1990; Harré, 1991; Harré and Gillett, 1994; Harré and Langenhove, 1991), the idea of the discursive production of the self and the mind is related to the idea of positioning. This is not an idea of 'structural positioning' but rather a metaphorical notion of agents' positioning in imaginary positions as points of view or perspectives, which are related to certain images, metaphors, storylines, and conceptual tools to understand reality, and which are constantly evolving in an unpredictable way in and through the constant undetermined flux of their contextual negotiation by the participants' interaction in conversation. Here, discursive practices denote the different institutionalized uses of language, in discourses (plural) that can compete with each other in reality-shaping. The self is variously positioned in the continuous movement among multiple (complementary, differentiating, or even contradictory) positions in different situations, largely being the 'product' of the various evolving discursive practices. The subject also understands others and their actions by positioning them in existing mutually – but also subjectively – understood unfolding narratives.

Moving through contradictory positions that are intertwined with contradictory discursive practices can lead to contradictory selves. However, it is frequently expected by others that in and through grammatical rules (Harré, 1991: 60–2; Harré and Langenhove, 1991: 400) of the consistent, unitary, and continuous self, we will try to remedy the discursively expressed contradictions, though tolerating or even celebrating contradictions is always a possibility in the repertoire of the construction of our storyline (Davies and Harré, 1990: 59). For Harré (1991) clearly negates the idea of the singularity and continuity of the core 'I' that unifies experience and generates a kind of coherence in and through the multiplicity of subject-positions. That is to say, the 'I' does not synthesize the various perspectives into one non-consistent storyline in view of the flux of events and speech acts. For Harré, there is no such entity; the singular 'I' is produced in discourse (*ibid.*: 60) as a grammatically created expectation and meaningful self-image.

Now, one could claim that the absence of a singular 'I' that assures the continuity and unification of experiencing the self/other divisions and relationality might imply the impossibility of self-reflection and intentionality. For the possibility of self-reflection requires a generic idea of the distantiating of an 'I' from a conception of 'Me' and 'Others' – which also ensures stable intentions.

However, we need to underline that for Davies and Harré (1990: 52), position-taking allows for the *possibility* of self-reflection in positioning oneself, that is, for the possibility of a 'choosing subject'. And in Harré and Gillett (1994), it is clear that the participation of the active subject in discourses assumes both agential intentionality (ibid.: 117–18) and self-reflection (ibid.: 127). As they characteristically state, 'The individual, whose intentions are structured by and emerge from the positions taken up in a social context and the discourses that pervade and structure it, is able to choose the rule-governed techniques that they will use to organize their psychological responses' (ibid.: 122). A careful reading of Harré and Gillett (ibid.) reveals their *tendency* to conceive of self-reflection – though they do not explicitly name it – and the agential ability (which is promoted in certain contexts; see ibid.: 127), as individuals are able to critically assess, negotiate, and modify alternative perspectives that are related to positions.

Regardless of the fact that Harré is not decisive about whether the idea of the discursive self can be combined with the ideas of self-reflection and intentionality, this is an excellent example of how thin the boundaries are between eliminative approaches and key, analytical distinctions. For, with reference to Figure 1, one could say that Harré (1991) follows (f) or even (a), while Harré and Gillett (1994) are clearly moving on towards (h) and downwards towards the final dilemma (Q4). Having noted this, it is time to ask whether reflectivity and intentionality can help us to clarify what we mean by the notion of human agency, or, to put it better:

Q3: Can human reflectivity (and intentionality) have a collective and/or individual impact on the transformation of social forms?

Many authors tacitly assume that if (d), then (h), that is to say, they assume that if subjectivities can reflect on inner states and outer micro-/macro-entities, and if they can have intentions that refer to entities or others at these levels of interaction, then they can formulate projects that can 'cause' sociocultural transformation.

However, it is not self-evident that self-reflective, intentional creatures can be *agents*. Take any random example from the many functionalist accounts in the long history of social theory and we can see that they simultaneously assumed voluntaristic notions of action. We can find 'milder' versions of this phenomenon in the contemporary literature in the tradition of critical realism, a school of thought that has clearly – and persuasively at some points – denounced functionalism. And yet, the idea of emergence has triggered interpretations in which agents are understood as both self-reflective and intentional and, at the same time, lower-level entities that participate in modes of circular causality in the composition of higher-level entities (for a discussion, see Bouzanis, 2016).

But we do not need to invoke naturalist imageries, like emergentist metaphors, to sustain the paradox. In a realist fashion, one could insist that social structures and

mechanisms are not malleable, that their causal powers persist and resist ideational elaboration, and that, even if social structures are activity-dependent, agents can reflect on their personal path and possibilities of action only through an objectivated, restrictive, and rigid world. Indeed, the idea of a 'rigid social structure' can entail low levels of upward social mobility or capacity of micro-contextual change. Adams (2006), for example, claims that we should distinguish between agents' capacity for reflection and their agential impact on and relative autonomy from structures of resources that still play a significant constraining role: self-reflection, for certain groups, 'in this context does not bring choice, just a painful awareness of the lack of it' (ibid.: 525). In this case, what we have is a picture of social actors with self-reflection but without agency, which seems to take us in the opposite direction to Bourdieu's intention to escape from the philosophy of the subject without doing away with the agent (Bourdieu and Wacquant, 1992: 121). We should note that, depending on how different authors understand intentionality, this strand can be placed in either (e) or (g) in Figure 1.

Caetano (2019: 156–8) is right that self-reflection does not necessarily imply action and/or social change, and that when self-reflection does lead to action or actions that aim at pursuing a personal project, they can unintentionally promote reproduction or partial transformation, or they can have unintentional (reproductive or transformational) consequences. Here, we need to distinguish between unintended results and unintentional consequences. The term *unintended results* designates our common experience that sociocultural and organizational changes are rarely the result of anyone's plans, which means that our plans and personal projects that do aim at institutional or organizational transformation are rarely fulfilled, at least in the way we have been dreaming of and planning them; social changes rarely correspond to the aspirations of one group, let alone of one single agent. This is the case not only because of the 'fact' that, in most cases, there are various interested groups and agents struggling for different or even conflicting outcomes, so that the final outcome is a product of negotiation and power play, but also because of the key role of *unintentional consequences*: broader side effects of the sociocultural interaction, which were not conceived of (or even imagined) as possible consequences in the context of the existing knowledge(s) that are discursively reproduced in the certain field of the power play.

Archer partly recognizes this dimension when stating that 'the unintended element largely results from group conflict and concession which altogether mean that the consequential elaboration is often what no-one sought or wanted' (Archer, 1995: 91, 143). This analysis leads us to the last of our four questions.

Q4: Does the idea of collective/individual impact on social forms allow for the possibility of intentional, synchronic, or diachronic reflective political praxis and self-governance?

It should by now be obvious to the reader that this article is taking the opposite approach to that which usually guides theoretical construction: instead of starting from a specific theory and assessing it in terms of whether it accounts for self-reflection and agency, we began by making some useful, generic distinctions in the first and second sections,

and now we turn the tables and follow an imaginary path (set out in Figure 1) of certain theoretical dilemmas that have arisen in relation to self-reflection. This allows us to identify common challenges within allegedly radically different approaches proposed in radically different traditions of thought.

If we wish to avoid (g) and admit that our intentional, reflected actions can have various kinds of effect on social forms, such as unintended results and unintentional consequences, we need to contemplate the ways in which our intentional acts can be either reproductive or transformational. Can we have an intentional impact on the organizational culture and/or the organizational structure? Can we *intentionally* change the rules of the games we participate in? And what would it mean to take path (j) instead of (i), which leads us to the idea that (collective) self-organization and self-government may be mere slogans, albeit illusionary ones.

Archer (2013) is right to criticize the parochial idea of self-organization in the recent neo-functional tendency in social theory to transfer imageries taken from the natural sciences – such as cybernetics, adaptive systems theory, and chaos and complexity theories – to the study of the social world. Yet, in her attack on this idea of systemic self-organizing, she also negates the possibility of individual and collective (ibid.: 161) self-governance:

In my trilogy of books on reflexivity, although I have indeed talked a lot about ‘people exercising some governance on their own lives’, this is far from considering agents to be paradigms of *self-government*. Granted that the majority of agents do have unimpaired powers of reflexivity, used in mental activities (such as planning, decision-making, budgeting, clarifying, self-monitoring), this does not make them self-governing. They are not because their autobiographies (singular or collective) are never made in circumstances of their own choosing. (ibid.: 162)

However, even though the sociocultural conditions are never of our own choosing, can we send the message, as social scientists, that social movements, interested elites and lobbies, political parties, corporate agents, and – why not? – *scientists* cannot have any kind of intentional effect on unequal structures of exploitation, institutional biases, and ideational frames of legitimacy and authority? Can we ever be not only agents, moving in and through objective, intention-independent conditions, but also participants in political praxis? If not, does this mean that dominant social theories of self-dialogue and/or self-reflection are not in harmony with commonsensical, everyday conceptions of democratic participation, conceptions that frequently assume *reflectivity, intentionality, agency, and effective participation in the body politic*, as I described in the first pages of this article.

Whipple (2005) draws an important comparison between (a) a classic debate between Lippmann’s elitist vision of democracy and Dewey’s participatory vision of democracy (in a similar manner to the above analysis); and (b) the contrast between the Habermasian notions of rationality and consensus and Dewey’s notion of reflective agency. Without implying that Habermas implicitly stands for elitist visions, Whipple (ibid.: 170) is right to point out that the resurrection of participatory visions of democracy goes hand in hand with the notion of critique, disagreement, and reflective agency, while approaches

of rationality and consensus are prone to eliminate what is essential to democratic dialogue – dissent and criticism – not as counterproductive forces but as the necessary elements of social interaction.

Whipple (2005) also provides an incisive reading of Dewey's efforts to formulate a political theory of agency by noting that he tried to combine reflectivity and habit. I would add to these remarks that even in elitist theories of democratic representation, self-reflection is both a necessary theoretical assumption and a tacit expectation. For even in elitist visions of political representation, processes such as decision-making in view of social conditions, crisis management, reactionary movements, conflicting roles, ideological battles in the parliament, voting itself as an intentional act, organizational power plays, conflicting institutional settings in civil society, and negotiations and compromises of various kinds are all socio-political phenomena in which self-reflection is not only assumed and expected but also a theoretical necessity. What is frequently neglected in this discussion is that negotiations and compromises – expressing the essence of unintended results – are products of reflection.

Therefore, while Whipple (2005) makes an intriguing remark on linking democratic participation, openness to scientific criticism, and human reflectivity – and contrasting these with elitist visions of democracy implicitly linked to ideals of rationality, technocratic expertise, and consensus – I would like to note that in fact, in the various possible visions of political representation, whether participatory or elitist, the protagonists are *assumed* to exercise time-consuming and interrogative (self-)reflective conduct, all the time.

Shall we summarize that our best social theories have downplayed self-reflection, and also that dominant contemporary theories of human reflectivity are not in harmony with commonsensical ideas of democratic participation? One could object to the phrasing of this question, in that the assumption of self-reflective subjectivity is an ontological one, and ontological assumptions do not necessarily need to comply with ordinary assumptions. Yet, even in the philosophy of science – which is focused mainly on the natural sciences – while it is frequently announced that our ontologies should be consistent with the findings of our best sciences (see e.g. Chakravarty, 2013: 37), it is also widely recognized that 'when constructing metaphysical models, ordinary experience has a privileged role in the sense that it provides a baseline or a starting point' (Paul, 2012: 16) and also that 'our ontology and ideology should explain the ordinary and scientific facts that we accept, and those we revise' (Ladyman, 2012: 41). In this case, how can the assumption of (the possibility of) self-reflective subjectivity, as related in greater sociological theoretical programmes, *not* be in tandem with our most basic everyday expectations that we hold from each other? And, more specifically, can our conceptualizations of self-reflection theoretically exclude political praxis? In what follows, I will take sides on the question of how we can organize our future efforts to understand human reflectivity.

Instead of a conclusion: Reinstating the ontological assumption of self-reflective subjectivity

For some sociologists, it would sound like a contradiction in terms: the paradox of joining the notion of reflectivity and the idea of subjectivity in the same expression, when for

certain sociological traditions subjectivity implies a passive internalization of intersubjective sets of norms, or even a merely epiphenomenal sense of the uniqueness of identity on the surface of social practices. Yet different authors define discourse in various ways: as an overarching veil of meaning setting the categorical and normative limits of local understandings; or as a mosaic of relatively independent islands of meaning, each a microcosm with its own norms and/or rules of normality; or as a fragmented surface of ever-changing embodied practices.

But there is an alternative notion of discourses, as various institutionalized, recursive sets of symbolic interactions in and through which participants draw on similar, complementary, conflicting, or even incommensurable world views, which thus relationally constitute the ideational background of these discourses. In this sense, discourses become the institutionalized medium of the modification of the ideational background as the intersubjectively shared set of world views on which agents meaningfully draw, so as to understand the existential possibilities of what is 'possible' or 'expected' in each micro-context.

Contra Strand and Lizardo (2015), we need to disentangle the ideational, institutional, and material levels of the social worlds so as to place the self-reflective participant as an agent who – in and through discourse – critically reflects on existing world-imageries and normative orientations, as well as institutional and material settings. This view of 'discourse and subjectivity' points to the idea that the self-reflective subjectivity is the prerequisite for participation in the institutionalized, recursive sets of interactions (discourses) in and through which agents *symbolically interact* with and make sense (fallibly) of world-imageries, normative orientations, roles and rules, and material structures and objects of everyday use. Symbolic battles over the 'appropriate' meaning of abstract concepts like democracy, equality, and justice indicate a view of discourse that assumes reflective agents rather than 'passive' script players.

It is unintelligible how it has become the norm to talk about 'discourses' as the imaginary 'space' of 'passive' internalization, without also allowing for a dialogic subjectivity that can reflect on the various levels we discussed in the first section. This notion of subjectivity highlights the idea of the internalization of world views and normative orientations, as one moment of the dialectic between internalization and agency (see Bouzanis, 2016). This approach can also ensure that the multivocality of the self does not negate the continuity and integrity of the subject (on this, see Falmagne, 2004). The internalization of multiple perspectives as related to various world views and normative orientations is in consistence with the idea of a subject who critically investigates, compares, and tests these in formulating hypotheses.

This article also highlights the lack of depth in our understanding of the modes of processes of self-reflection, emphasizing the fact that while Wiley's 'I'–'Me'–'You' model is one of the more sophisticated that we have in our explanatory repertoire, we still need further interdisciplinary collaboration in the human sciences, so as to achieve a holistic view of how self-reflection can be theoretically combined with institutional routines, habits, and embodied dispositions. What is required is a transdisciplinary platform model allowing that the self-reflective subjectivity can critically *objectify* various micro-contexts and macro levels of discursive participation. And as I have explained, this call for key distinctions, instead of the horizontal conflation entailed in the imagery of the

‘passive’ subjectivity, does not necessarily lead us to a realist, Archerian ontology involving an analytical distinction between the first-person ontology of the subject and the third-person ontology of objective social structures.⁸

The elimination of key distinctions in social theory can be helpful in escaping from classic socio-philosophical puzzles (such as how to combine the various levels of analysis) that now might appear irresolvable, or unnecessary, and can also offer us a simple model of (inter)action in micro-contexts that can be projected in various sub-worlds (from primitive cultures to contemporary laboratories). But this sounds more like the trumpet of retreat into simplicity as the sole epistemic virtue, rather than a careful, continuous effort to *reflect* on micro/macro levels of interaction – levels of participation in which self-reflective subjects are *expected to behave as* intentional and accountable agents.


Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Christoforos Bouzanis  <https://orcid.org/0000-0003-2424-0124>

Notes

I would like to thank Michael Vaughan for his comments on several versions of this article. I am also grateful to Stephen Kemp for the many thought-provoking discussions that we had in Edinburgh on some of the topics examined in this article.

1. For a valuable overview, see Smith (2005).
2. Note that Archer (2003, 2007, 2012) offers a promising categorization of species of reflexivity – consisting of communicative reflexives, autonomous reflexives, meta-reflexives, and fractured reflexives – that can be utilized in social research on self-reflection.
3. On this attack against philosophical ‘representationalism’, see also Christopher and Bickhard (2007) and Piirainen (2018).
4. Note that Lizardo (2017) offers an interesting typology of culture, a key component of which is the distinction between declarative culture, which is open to inspection via ‘reflective cognitive acts’ (ibid.: 92), and non-declarative culture, which is non-reflective and intention-independent (ibid.: 93). It is hard to see how one can combine this interesting distinction with Lizardo’s analysis of habit, in which habitus is the main cognitive resource for action, even if we set aside that his additional distinction between personal culture and public culture confuses the discussion about what culture is – for in the literature, a well-established assumption about culture is that it consists of shared ideational elements or interpersonal attributes.
5. Teo (2017) is right to claim that we need an integrated theory of subjectivity that can stand as a key element in various social and humanistic studies. I would add that to be consistent with the ideal of transdisciplinarity itself, and also with social scientific practices that aim to reflect on

- human conduct, such an integrated theory can be conceived of only as a framework that integrates the notion of subjectivity with the assumption of self-reflection.
6. For an interesting comparison between these two key authors, see Sullivan (2010).
 7. The multivocality of the self can imply a state of multi-intentionality of the subject, who holds different intentions at different levels of reflection. Contra certain authors who have recently argued for notions of collective intentionality (Tomasello, 2020; Tuomela, 2013), we need to remind ourselves that in social theory we have taken important steps towards understanding that agents are structurally positioned, having conflicting interests even within the same organization, and that they share multiple perspectives – to the extent that cultural integration is a myth (see Archer, 1985). Therefore, it is a logical trap to move from the idea of shared ideational backgrounds (even those consisting of similar or complementary world-imageries) to simplistic notions of collective intentions.
 8. Aside from the above-mentioned complicity with part of Harré's work, there are sophisticated constructionist views in which discursive participation is combined with self-reflection and self-dialogue (see e.g. Burns and Engdahl, 1998a, 1998b).

References

- Adams, M. (2006) 'Hybridizing Habitus and Reflexivity: Towards an Understanding of Contemporary Identity?', *Sociology* 40(3): 511–28.
- Alvesson, M. and Kärreman, D. (2011) 'Decolonizing Discourse: Critical Reflections on Organizational Discourse Analysis', *Human Relations* 64(9): 1121–46.
- Archer, M. (1985) 'The Myth of Cultural Integration', *British Journal of Sociology* 36(3): 333–53.
- Archer, M. (1995) *Realist Social Theory: The Morphogenetic Approach*. Cambridge: Cambridge University Press.
- Archer, M. S. (1996) *Culture and Agency: The Place of Culture in Social Theory*. Cambridge: Cambridge University Press.
- Archer, M. (2000) *Being Human: The Problem of Agency*. Cambridge: Cambridge University Press.
- Archer, M. (2003) *Structure, Agency and the Internal Conversation*. Cambridge: Cambridge University Press.
- Archer, M. (2007) *Making Our Way Through the World: Human Reflexivity and Social Mobility*. Cambridge: Cambridge University Press.
- Archer, M. (2010a) 'Can Reflexivity and Habitus Work in Tandem?', in M. Archer (ed.) *Conversations About Reflexivity*. London: Routledge, pp. 123–43.
- Archer, M. (2010b) 'Routine, Reflexivity, and Realism', *Sociological Theory* 28(3): 272–303.
- Archer, M. (2012) *The Reflexive Imperative in Late Modernity*. Cambridge: Cambridge University Press.
- Archer, M. (2013) 'Morphogenic Society: Self-Government and Self-Organization as Misleading Metaphors', in M. Archer (ed.) *Social Morphogenesis*. Dordrecht: Springer, pp. 145–64.
- Ashmore, M. (1989) *The Reflexive Thesis: Writing Sociology of Scientific Knowledge*. London: University of Chicago Press.
- Beck, U., Giddens, A., and Lash, S. (1994) *Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order*. Stanford, CA: Stanford University Press.
- Benbassat, N. (2020) 'Reflective Function: A Move to the Level of Concern', *Theory & Psychology* 30(5): 657–73.

- Blum, A. and McHugh, P. (1984) *Self-Reflection in the Arts and Sciences*. Atlantic Highlands, NJ: Humanities Press.
- Bourdieu, P. and Wacquant, L. J. (1992) *An Invitation to Reflexive Sociology*. Cambridge: Polity Press.
- Bouzanis, C. (2016) 'Ontogenesis Versus Morphogenesis Towards an Anti-realist Model of the Constitution of Society', *Human Studies* 39(4): 569–99.
- Bouzanis, C. (2017) 'For Reflexivity as an Epistemic Criterion of Ontological Coherence and Virtuous Social Theorizing', *History of the Human Sciences* 30(5): 125–46.
- Bouzanis, C. and Kemp, S. (2020) 'The Two Stories of the Habitus/Structure Relation and the Riddle of Reflexivity: A Meta-theoretical Reappraisal', *Journal for the Theory of Social Behaviour* 50(1): 64–83.
- Burkitt, I. (2012) 'Emotional Reflexivity: Feeling, Emotion and Imagination in Reflexive Dialogues', *Sociology* 46(3): 458–72.
- Burns, T. R. and Engdahl, E. (1998a) 'The Social Construction of Consciousness Part 1: Collective Consciousness and Its Socio-cultural Foundations', *Journal of Consciousness Studies* 5(1): 67–85.
- Burns, T. R. and Engdahl, E. (1998b) 'The Social Construction of Consciousness Part 2: Individual Selves, Self Awareness, and Reflectivity', *Journal of Consciousness Studies* 5(2): 166–84.
- Caetano, A. (2019) 'Designing Social Action: The Impact of Reflexivity on Practice', *Journal for the Theory of Social Behaviour* 49(2): 146–60.
- Caldwell, R. (2007) 'Agency and Change: Re-evaluating Foucault's Legacy', *Organization* 14(6): 769–91.
- Chakravarty, A. (2013) 'On the Prospects of Naturalized Metaphysics', in D. Ross, J. Ladyman, and H. Kincaid (eds) *Scientific Metaphysics*. Oxford: Oxford University Press, pp. 27–50.
- Christopher, J. C. and Bickhard, M. H. (2007) 'Culture, Self and Identity: Interactivist Contributions to a Metatheory for Cultural Psychology', *Culture & Psychology* 13(3): 259–95.
- Crossley, N. (2003) 'From Reproduction to Transformation: Social Movement Fields and the Radical Habitus', *Theory, Culture & Society* 20(6): 43–68.
- Davies, B. and Harré, R. (1990) 'Positioning: The Discursive Production of Selves', *Journal for the Theory of Social Behaviour* 20(1): 43–63.
- Decoteau, C. L. (2016) 'The Reflexive Habitus: Critical Realist and Bourdieusian Social Action', *European Journal of Social Theory* 19(3): 303–21.
- Elder-Vass, D. (2007) 'Reconciling Archer and Bourdieu in an Emergentist Theory of Action', *Sociological Theory* 25(4): 325–46.
- Elder-Vass, D. (2010) *The Causal Power of Social Structures: Emergence, Structure and Agency*. Cambridge: Cambridge University Press.
- Emirbayer, M. and Mische, A. (1998) 'What Is Agency?', *American Journal of Sociology* 103(4): 962–1023.
- Falmagne, R. J. (2004) 'On the Constitution of "Self" and "Mind": The Dialectic of the System and the Person', *Theory and Psychology* 14(6): 823–46.
- Farrugia, D. and Woodman, D. (2015) 'Ultimate Concerns in Late Modernity: Archer, Bourdieu and Reflexivity', *British Journal of Sociology* 66(4): 626–44.
- Gergen, K. J. (1985) 'The Social Constructionist Movement in Modern Psychology', *American Psychologist* 40(3): 266–75.

- Gergen, K. J. (2009) *Relational Being: Beyond Self and Community*. Oxford: Oxford University Press.
- Gergen, K. J. (2011) 'The Self as Social Construction', *Psychological Studies* 56(1): 108–16.
- Giddens, A. (1984) *The Constitution of Society: Outline of the Theory of Structuration*. Berkeley: University of California Press.
- Gillespie, A., ed. (2006) *Becoming Other: From Social Interaction to Self-Reflection*. Greenwich, CT: IAP.
- González Rey, F. L. (2019a) 'Subjectivity and Discourse: Complementary Topics for a Critical Psychology', *Culture & Psychology* 25(2): 178–94.
- Gonzalez Rey, F. (2019b) 'Subjectivity in Debate: Some Reconstructed Philosophical Premises to Advance Its Discussion in Psychology', *Journal for the Theory of Social Behaviour* 49(2): 212–34.
- Habermas, J. (1996) *Theory and Practice*, trans. J. Viertel. Cambridge: Polity Press.
- Habermas, J. (2007) 'The Language Game of Responsible Agency and the Problem of Free Will: How Can Epistemic Dualism Be Reconciled With Ontological Monism?', *Philosophical Explorations* 10(1): 13–50.
- Harré, H. R. (1975) 'Images of the World and Societal Icons', in H. G. Zilian, K. Knorr-Cetina, H. Strasser, and K. D. Knorr (eds) *Determinants and Controls of Scientific Development*. Dordrecht: Springer, pp. 257–83.
- Harré, R. (1991) 'The Discursive Production of Selves', *Theory & Psychology* 1(1): 51–63.
- Harré, R. and Gillett, G. (1994) *The Discursive Mind*. Thousand Oaks, CA: Sage.
- Harré, R. and Langenhove, L. V. (1991) 'Varieties of Positioning', *Journal for the Theory of Social Behaviour* 21(4): 393–407.
- Hermans, H. J. M. (2001) 'The Dialogical Self: Toward a Theory of Personal and Cultural Positioning', *Culture and Psychology* 7(3): 243–81.
- Hermans, H. J. (2002) 'The Dialogical Self as a Society of Mind: Introduction', *Theory & Psychology* 12(2): 147–60.
- Hermans, H. J. (2003) 'The Construction and Reconstruction of a Dialogical Self', *Journal of Constructivist Psychology* 16(2): 89–130.
- Hermans, H. J. (2006) 'The Self as a Theater of Voices: Disorganization and Reorganization of a Position Repertoire', *Journal of Constructivist Psychology* 19(2): 147–69.
- Hermans, H. J. (2012) 'Dialogical Self Theory and the Increasing Multiplicity of I-positions in a Globalizing Society: An Introduction', *New Directions for Child and Adolescent Development* 137: 1–21.
- Holmes, M. (2010) 'The Emotionalization of Reflexivity', *Sociology* 44(1): 139–54.
- Joas, H. (1996) *The Creativity of Action*. Cambridge: Polity Press.
- Kemp, S. (2012) 'Interests and Structure in Dualist Social Theory: A Critical Appraisal of Archer's Theoretical and Empirical Arguments', *Philosophy of the Social Sciences* 42(4): 489–510.
- Ladyman, J. (2012) 'Science, Metaphysics and Method', *Philosophical Studies* 160(1): 31–51.
- Lizardo, O. (2004) 'The Cognitive Origins of Bourdieu's Habitus', *Journal for the Theory of Social Behaviour* 34(4): 375–401.
- Lizardo, O. (2017) 'Improving Cultural Analysis: Considering Personal Culture in Its Declarative and Nondeclarative Modes', *American Sociological Review* 82(1): 88–115.

- Lizardo, O. (2021) 'Habit and the Explanation of Action', *Journal for the Theory of Social Behaviour*, advance online publication, available at: <https://doi.org/10.1111/jtsb.12273>
- Martin, J. and Gillespie, A. (2010) 'A Neo-Meadian Approach to Human Agency: Relating the Social and the Psychological in the Ontogenesis of Perspective-Coordinating Persons', *Integrative Psychological and Behavioral Science* 44(3): 252–72.
- Mutch, A. (2004) 'Constraints on the Internal Conversation: Margaret Archer and the Structural Shaping of Thought', *Journal for the Theory of Social Behaviour* 34(4): 429–45.
- Paul, L. A. (2012) 'Metaphysics as Modeling: The Handmaiden's Tale', *Philosophical Studies* 160(1): 1–29.
- Pels, D. (2000) 'Reflexivity: One Step Up', *Theory, Culture & Society* 17(3): 1–25.
- Piironen, T. (2018) 'A Meaning Holistic (Dis)solution of Subject–Object Dualism – Its Implications for the Human Sciences', *History of the Human Sciences* 31(3): 64–82.
- Pollner, M. (1991) 'Left of Ethnomethodology: The Rise and Decline of Radical Reflexivity', *American Sociological Review* 56(3): 370–80.
- Popper, K. R. (1972) *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press.
- Porpora, D. V. and Shumar, W. (2010) 'Self Talk and Self Reflection', in M. Archer (ed.) *Conversations About Reflexivity*. London: Routledge, pp. 206–20.
- Rafieian, S. and Davis, H. (2016) 'Dissociation, Reflexivity and Habitus', *European Journal of Social Theory* 19(4): 556–73.
- Smith, R. (2002) 'Self-Reflection and the Self', in R. Porter (ed.) *Rewriting the Self: Histories From the Middle Ages to the Present*. London: Routledge, pp. 61–70.
- Smith, R. (2005) 'Does Reflexivity Separate the Human Sciences From the Natural Sciences?', *History of the Human Sciences* 18(4): 1–25.
- Sosa, E. (2009) *Reflective Knowledge: Apt Belief and Reflective Knowledge: Vol. 2*. Oxford: Oxford University Press.
- Strand, M. (2019) 'Practice Theory and Conservative Thought', *History of the Human Sciences* 32(5): 108–34.
- Strand, M. and Lizardo, O. (2015) 'Beyond World Images: Belief as Embodied Action in the World', *Sociological Theory* 33(1): 44–70.
- Strand, M. and Lizardo, O. (2017) 'The Hysteresis Effect: Theorizing Mismatch in Action', *Journal for the Theory of Social Behaviour* 47(2): 164–94.
- Sullivan, P. (2010) 'Vygotskian Dialectics and Bakhtinian Dialogics: Consciousness Between the Authoritative and the Carnavalesque', *Theory & Psychology* 20(3): 362–78.
- Sweetman, P. (2003) 'Twenty-First Century Dis-ease? Habitual Reflexivity or the Reflexive Habitus', *Sociological Review* 51(4): 528–49.
- Tauber, A. I. (2005) 'The Reflexive Project: Reconstructing the Moral Agent', *History of the Human Sciences* 18(4): 49–75.
- Taylor, C. (1995) 'The Dialogical Self', in R. F. Goodman and W. R. Fisher (eds) *Rethinking Knowledge: Reflections Across the Disciplines*. Albany, NY: State University of New York Press, pp. 57–68.
- Teo, T. (2017) 'From Psychological Science to the Psychological Humanities: Building a General Theory of Subjectivity', *Review of General Psychology* 21(4): 281–91.
- Tomasello, M. (2020) 'The Role of Roles in Uniquely Human Cognition and Sociality', *Journal for the Theory of Social Behaviour* 50(1): 2–19.

- Tuomela, R. (2013) *Social Ontology: Collective Intentionality and Group Agents*. Oxford: Oxford University Press.
- Von Wright, J. (1992) 'Reflections on Reflection', *Learning and Instruction* 2(1): 59–68.
- Whipple, M. (2005) 'The Dewey-Lippmann Debate Today: Communication Distortions, Reflective Agency, and Participatory Democracy', *Sociological Theory* 23(2): 156–78.
- Wiley, N. (1994) *The Semiotic Self*. Cambridge: Polity Press.
- Wiley, N. (2006a) 'Inner Speech as a Language: A Saussurean Inquiry', *Journal for the Theory of Social Behaviour* 36(3): 319–41.
- Wiley, N. (2006b) 'Pragmatism and the Dialogical Self', *International Journal for Dialogical Science* 1(1): 5–21.
- Wiley, N. (2010) 'Inner Speech and Agency', in M. Archer (ed.) *Conversations About Reflexivity*. London: Routledge, pp. 17–38.
- Woolgar, S. (1988) 'Reflexivity Is the Ethnographer of the Text', in S. Woolgar (ed.) *Knowledge and Reflexivity: New Frontiers in the Sociology of Knowledge*. London: SAGE, pp. 14–34.
- Woolgar, S. and Ashmore, M. (1988) 'The Next Step: An Introduction to the Reflexive Project', in S. Woolgar (ed.) *Knowledge and Reflexivity: New Frontiers in the Sociology of Knowledge*. London: SAGE, pp. 1–11.
- Wrong, D. H. (1961) 'The Oversocialized Conception of Man in Modern Sociology', *American Sociological Review* 26: 183–93.

Author biography

Christoforos Bouzanis is a lecturer in people and organizations at Swansea University, UK. His main research interest lies in the interrelations between philosophy and the social sciences. He is currently working on his project of defining socio-theoretical anti-realism as an alternative path for the sociology of knowledge, by critically drawing on the traditions of metaphysical realism, critical realism, social constructionism, and pragmatism.