

Institutional Conflicts of Interest: A Novel Functionalist Account

Armin Schulz

Department of Philosophy - University of Kansas

Lawrence, KS 66045, USA

awschulz@ku.edu

Abstract

Much of economics is dedicated to studying conflicts of interest. What is less well studied is the question of how to make sense of conflicts of interest involving non-individuals like social institutions or social collectives. However, the latter is very important, too: from the occurrence of institutional corruption to the creation of workplaces with much internal conflict, understanding clashes between the interests of individuals and non-individuals is an important social phenomenon. How, though, can this phenomenon be studied? What, exactly, are the “interests” of social institutions or collectives? To make progress with this, this paper uses recent work in social functionalism to develop a theory of cooperation and conflict involving non-individualist social entities. To make the discussion precise, the paper focuses on principal / agent problems (though it is not restricted to the latter). The paper ends by applying this theory to cases of internal conflicts in social collectives.

Acknowledgement: Work for this paper was graciously supported by Grant ID 63320 from the John Templeton Foundation. The author would like to thank the members of the grant team, the anonymous referees for this journal, and audiences at the Universities of Buffalo and Kansas for much helpful discussion.

Institutional Conflicts of Interest: A Novel Functionalist Account

I. Introduction

Social cooperation and coordination have been key topics of discussion in the social sciences since the days of Adam Smith at least. In turn, a central concern at the heart of these discussions are conflicts of interest (COI): different individual agents are assumed to have interests that are not fully aligned, and they thus face the problem of whether and how to cooperate and coordinate their actions.

However, as this paper shows, this purely individualistic view of COI overlooks an important version of this kind of problem: that involving at least some *non-individualist* social entities. It is not just the case that individual economic agents need to balance their interests *against each other*—they also need to do so against non-individualist entities, such as social institutions, norms, or social collectives (on the metaphysics of social institutions, see Searle, 2005, Searle, 1995, Guala, 2016). In these cases, the issue is not that different individuals disagree with each other; rather, the issue is that the individuals disagree with an entity—a firm or social institution, say—*as a whole*. In what follows, such entities are called “non-individualist social entities” (NISE). In short: COI seem to be able to also occur with NISE.

This is important to note, as it makes the existing, individualistic treatments of the problem of COI inapplicable. With NISE, it is not so easy to see what their interests could be—hence, it is not so easy to see how they could disagree with an individual. Now, this may be the right conclusion to draw: perhaps the appearance of the existence of COI involving NISE is just that—an appearance. However, as the rest of this paper shows, this conclusion is unwarranted: there is a compelling, functionalist way to make sense of NISE interests. One theoretical virtue of this

account is that it makes for simple analysis of COIs between individuals and NISEs by using existing COI frameworks. In this way, it can be shown that across a number of important domains (e.g. the corporation, democracy), we already have analytical tools in our hands that can make sense of these COIs and do not need a new one. In turn this allows us to grasp important theoretical and empirical phenomena that we would otherwise miss.

The paper is structured as follows. Section II makes the problem to be solved clearer: it is not obvious how traditional approaches towards COI will be able to deal with COI in NISE without an account of the interests of NISE. Section III develops a new account of this phenomenon by showing how “presentist social functionalism”—a theory which allows the assignment of functions to social institutions—can be used to assign interests to at least some NISE. The section goes on to show how this insight can be used to generalize the principal / agent framework to allow for non-individualistic COI. Section IV presents to a key application of this new account of COI: an expanded taxonomy of internal conflicts in economic collectives. Section V concludes.

II. Conflicts of Interest, the Principal / Agent Problem, and the Need to Go Beyond Individualism

Much work in economics (and the social sciences in general) seeks to make clear when, why, and how entities can balance their respective interests to achieve outcomes that they could not achieve on their own. Now, it is possible that these respective interests are aligned in such a way that each pursuing their own wants yields outcomes that are mutually consistent.¹ However,

¹ Famously, it is possible that divergent interests still yield socially beneficially outcomes: “[H]e intends only his own gain; and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. Nor is it always the worse for the society that it was no part of it. By pursuing his own interest, he frequently promotes that of the society more effectually than when he really intends to promote it.” (Smith, 1776,

while in principle possible, such cases of full interest alignment are, at best, benchmarks. In most realistic cases, a divergence exists: the goals and motivations of one individual in general do not coalesce with those of others, and conflicts thus arise. This is the problem of COI: how to understand and deal with the frequent occurrence of situations where what one party aims to achieve is not fully in alignment with the goals of other parties in the society in question.

Given this very abstract statement of the problem, it is unsurprising that many different approaches have been developed to understand the many different instances of the problem of COI—from social dilemmas (Skyrms, 2004) and corporate governance (Demski, 2003) to general equilibrium models (Mas-Colell et al., 1995) and legal theory (Fama and Jensen, 1983, Buchen, 2024). For concreteness, the focus in what follows will be on “principal / agent problems” (PAPs) (itself a specific instance of the large class of social dilemmas) (Ross, 1973, Grossman and Hart, 1983, Hart and Holmström, 1987, Skyrms, 2004). However, it is important to note from the get-go that this focus is not meant to be restrictive: much of what follows can be straightforwardly extended to other models of COI. This will also be made clearer below and in section IV.

PAPs are situations where one individual—the principal—asks another individual—the “agent”—to act on their behalf. Importantly, in these cases, there can and typically are misaligned interests. There is a lot of work, both formal and empirical, exploring this framework further. Many of these details are not central here, but it is helpful to consider briefly the outlines of the work / shirk game as a simple but telling example of this kind of framework (see also Glimcher et al., 2005). In this game, employees can choose to work (W) or shirk ($\sim W$); managers

book IV, chap. II). However, this is a slightly different point, as it concerns what is *socially* beneficial. The issue with COI, though, is whether the outcome is beneficial for the *parties in question*. As is well-known, the fact that a competitive market yields Pareto efficient outcomes does not mean that everyone gets all they want. That said, the notion of “social benefit” will become important again below.

can choose to inspect the work of the employee (I) or trust the work (~I). Working is costly for the employee, and they would prefer to shirk—but only if not inspected. Inspecting is costly for the manager, and they would prefer to trust the employee—but only if they are working:

Employee / Manager	I	~I
W	1, 2	1, 3
~W	0, -1	2, -2

[Figure 1: work / shirk game]

This game has a mixed strategy equilibrium, where workers work with probability $p = \frac{1}{2}$ and managers inspect with probability $q = \frac{1}{2}$. Note that this mixed equilibrium yields strictly less utility than what the employee would have gotten if they had got their way (shirk and not get caught) and what managers would have gotten if they had got their way (trust and have the employee work). For present purposes, there are three key points to note about principal / agents problems like the one in figure 1.

First, while, as just noted, the PAP is just one specific approach to the problem of COI, it still needs to be recognized as a central approach towards studying COI. In particular, it theoretically isolates a paradigmatic form of COI. In PAPs, it is presumed that the parties in question need to work together: that is why there is an agent that acts on behalf of the principal in the first place. The trouble, though, is that this need for cooperation coexists with divergent interests of the parties in question. Hence, these parties (continuously) need to figure out how to balance their respective interests with the need for continued cooperation and coordination. In this way, PAPs get to the heart of what the problem of COI is like. This also explains why PAPs are good

models for many empirical examples of COI. For example, in a firm, the owners and employees typically want different things (working vs. shirking, or doing work of one kind rather than another kind) or believe different things (how hard the agent work works, or what work is most important to be done now). This is exactly the nature of PAPs. The same, though, goes for many contractual relationship even outside of firms: if I contract you to cut my hair, you are acting on my behalf. This can create COI—e.g. if you are interested in displaying your skill as a hairdresser, but I am interested in getting a very simple cut only. In this way, many aspects of the economy can be approached by focusing on PAPs.

Second, though, much of what follows can be generalized to other models of COI. The point in what follows is to understand what the interests are of a non-individual are, so as to be able to provide a broader theory of COI. The focus on PAPs specifically is useful here, as it allows the discussion to have a depth that it would not otherwise have. However, much of what follows is easily translated into other types of situations involving COI. Indeed, section IV places PAPs into a broader typology of other approaches towards COI.

Third, PAPs as they are commonly understood are thoroughly individualist, in the sense that they concern the relationship between the interests of different *individual people*. As noted in the introduction, though, this overlooks an important set of COI: disagreements involving non-individual social entities (NISE). At times, an individual's actions are constrained, not just by what other individuals want, but also by the wants of firms, government agencies, and social institutions, treated as such. Indeed, it seems there can be COI purely between different NISE—different branches of government, for example, can be constitutionally set up to have differing “wants,” and the same goes for different parts of firms. In what follows, such COI involving

NISE will be referred to as “NICOI” (non-individualistic conflicts of interest).² Section IV will provide some further details of these cases, but for now, it is sufficient to note that, for both theoretical and empirical reasons, being able to study NICOI matters.

Theoretically, it is important to have a toolkit that is not *intrinsically* limited to individualist methodologies. Of course, the debate surrounding “microfoundations” in economics is old and complex (Little, 1998, Ross, 2005, Hoover, 2009, Hoover, 2010, Watkins, 1952, Elster, 1982, Jones, 1996, Hodgson, 2007, Langlois, 1985). However what matters here is just that it has become clear that it is implausible to think that there is a monistic resolution to this debate: individualistic approaches are plausible in some, but not all cases (Epstein, 2015, Ruiz and Schulz, 2023). Importantly, there is no compelling a priori reason to think that all internal conflicts fall into the former category: it is at least *prima facie* possible that there are internal conflicts between individuals and the larger social collective they are part of. Hence, it is important to provide a compelling theoretical grounding for instances of internal conflicts that are not purely individualistic in nature. This is especially so, since many prior treatments of these issues seem to be based only on imprecise blanket statements concerning the interests of, say “the establishment.” What the rest of the paper is trying to make clear, though is that it is possible to make compelling, precise, and formalizable sense NICOI.³

² This paper leaves it open whether the non-individualist entity is a collective agent of sorts. This may be plausible for some cases and not others, and, at any rate, these issues can be left open here. See also section IV below. In the same way, this paper does not consider the question of whether a NISE is an individual in the metaphysical sense—i.e. a particular, not a group or kind. The issue here is whether there can be COI where one of the parties is not an individual *person* (though they may still be an individual in the metaphysical sense of not being a kind).

³ Of course, this is not to say that all COI *featuring* social institutions need to be NICOI. There can be many cases where social institutions constrain people’s actions without these institutions themselves being a party to the conflict. For example, in several cultures, having a child without being married has or had serious consequences for the parents. In these cases, unmarried parents play a game not against the institution of marriage but against other people (the norm not to have children out of wedlock has an impact on the payoffs, but it is not a player). Still, the point in the text is just that not *all* COI should be assumed to be well handled in these ways. See also note 5.

Indeed, empirically, too, capturing NICOI matters, as they seem in fact quite common. The introduction has provided some examples, and sections III and IV will return to these issues as well, but a particularly prominent case concerns the “paradox of democracy” or the “paradox of liberty” (Popper, 1945). When all the people of a democracy agree that they no longer want to be self-governed, there is a conflict between all the people and the institution of democracy—not among individual people.⁴ (We may be tempted to say this is a conflict between current members of the society and potential future people—but since the latter do not actually exist, they are not well seen as individuals in their own right.)⁵ Importantly, as the below makes clearer, this is just the tip of the iceberg; once this sort of phenomena is recognized, many more instances of it become visible—with a clear instance being the case of a large, established corporation, as made clearer in section III below.

The rest of this paper will therefore fill this theoretical and empirical lacuna by developing an account that allows for NICOI. Importantly, as the rest of the paper also shows, this account can be made formally specific and integrated with the existing work on COI.⁶

⁴ The paradox of democracy is sometimes understood as a kind of cognitive contradiction: democracy entails majority rule, but a majority could vote for tyranny. In this paper, the term is used more broadly to refer to the idea that a democracy cannot persist when the people vote to abolish voting, even though voting is at the heart of democracy. Hence, the will of the voters can be in conflict with the survival interests of the system as a whole. Importantly also, there can be a strategies a democracy can use to protect itself from being overturned by the will of the voters, such as ringfencing the ability of voters to change the voting procedures at the heart of the democracy. This will be made clearer in the discussion in concerning the corporation in sections III and IV. See also note 10.

⁵ This point also addresses a different version of an institutional account of COI’s: we might understand all COI’s as concerning individual people; it is just that, sometimes, these COI’s are mediated by institutional structures. This might still be a non-individualism of sort—we might not be able to reduce the institutional roles to individual thoughts and actions (though that is debatable)—but it would not see the institution as a party to the COI. However, in the above cases, this is not what is going on (though it may of course apply in other cases): the conflict is not between individual people at all—again keeping in mind the fact that merely possible future people are not actually people with interests at all. I thank an anonymous referee for useful discussion of this issue.

⁶ Another approach towards firms, specifically, is based on behavioral routines (Nelson and Winter, 1982, Nelson and Winter, 2002, Nelson et al., 2018). This, though, could be made fit into any of the above approaches, and thus will not be discussed further here. The same goes for the equilibrium view of institutions (Schotter, 1981, Aoki, 2001, Greif and Kingston, 2011, see also Guala, 2016), as it does not, itself, provide a way to model the conflict *between* institutions and individuals. Of course, the latter is able explain how the actions and thoughts of individual people influence the institutions they are part of. However, here the issue is a different one: whether institutions can

III. A New Approach

Do NISEs have interests? This is a long-debated topic among philosophers and social scientists alike. In order to extend PAPs (and other kinds of COI) to cases of NICOI, we would need to be able answer this question in the affirmative.

Now, on the face of it, it may be thought that there are *individualistic* ways to do just that. While it may be true that not all COI are among individuals, it might be thought that, to the extent that they are not, individuals are still the basis on which all COI rest: non-individualistic interests may be thought to be derivable from individual interests. For example, one could follow a broadly social choice-theoretic, aggregating framework, and build up the interests of the non-individuals social entities from that of their members (List, 2012, List and Pettit, 2011). We could say that the interest of a corporation is the average of the interests of its members. Or we could follow some classic political theory and say that the interest of a *state* is the majority of the interests of its citizens. Or we could (very loosely following Rawls) have it that the interest of a collective is the interest of the least advantaged of its members. Many other approaches exist as well.

However, this individualistic, social choice theoretic approach will, *in general*, not be adequate. First, as the history of social-political thought makes clear, it is not obvious what the right way is of aggregating interests. Since aggregation is the very nature of the individualistic accounts, this lack of agreement on the right aggregation function means that there is a major lacuna at the heart of these accounts. Second and relatedly, it is the core insight of social choice theory that, depending on the details of the aggregation function used, many seemingly possible

be seen to have interests of their own. This question can be asked independently of whether institutions are seen as rules, equilibria, or something in between (Guala, 2016).

collective interest assignments may turn out to, in fact, be inadmissible (Arrow, 1951, Sen, 1970, List, 2012).

Now, these issues may be solvable: perhaps a suitable aggregation function can be found that can be defended on both social-political and social-choice-theoretic grounds. However, even this would not fundamentally alter the issues here, as there is a third problem with this individualistic aggregating approach: not all NISE are collectives with a contemporaneous membership. For example, it is not the case that individual people are *parts* social institutions like democracy (Ruben, 1985). People can choose to follow or abide by these institutions, and these institutions depend on the fact that there are people with interests in society (Searle, 1995, Searle, 2005)—but people do not constitute them, in the sense that Ms. Smith’s 3rd Grade Class of 2025 is constituted by all and only the students in it. Unlike spatiotemporally circumscribed entities like Ms. Smith’s 3rd Grade Class of 2025, social institutions like the US democratic system extend beyond the current members of a society, and also include potential future people. Importantly, also, we cannot take the interests of these potential future people into account like we can those of actual people: potential future people do not actually exist, and thus do not have interests that could be aggregated over. As noted earlier, this can be seen as a version of the “paradox of democracy:” not everything “the people” want is in line with what is in the interest of democracy. Because of this, the individualistic, social-choice theoretic approach is too constrained to be generally compelling.

However, how can a non-reductive, non-individualistic treatment of the interests of a social non-individualist entity be determined? In response, the rest of this section will argue that these interests are best seen to lie in the *function* of the NISE: what this entity is *for*. This is a good place to start looking for these interests, as a social entity’s *purpose* sets out, in some sense at

least, what it would be good for the entity (whether individual or not) to be or do. The entity's purpose describes what it ought to be or do—and being and doing what it ought to be or do is, in this general way, in its interest. To appreciate this better, two points need to be noted.

On the one hand, some account needs to be provided that explains exactly what the purpose of a NISE is. Otherwise, this would be a matter of replacing one mysterious notion—that of the *interest* of a NISE—for another—that of the *purpose* of a NISE. On the other hand and somewhat relatedly, some reason needs to be provided for thinking that many NISE's *have* functions (so understood). If this were not so, we would need to conclude that many NISE's do not have interests either.

To address these points, the rest of this paper relies on a specific, recently developed account of the function (or purpose) of social institutions: Presentist Social Functionalism (PSF). To be sure, other accounts of social functionalism exist, too (Merton, 1968, Parsons, 1951, Pettit, 1996), and much—though not necessarily all—of what follows could be reformulated with these other accounts in the background as well. The focus on PSF is helpful, though, as it (a) avoids some of the key pitfalls that have been raised concerning social functionalism in general (for discussion, see e.g. Elster, 1982, Langlois, 1985, Schulz, 2025, Pettit, 1996), and (b) this account has some important features that makes it a particularly compelling contender in the present context (Schulz, 2022). Indeed, as the rest of this section makes clear, this account is well placed to be extended so as to ground the interests of NISE's.

At the heart of the presentist social functionalist account is the idea that feature F of social entity E makes for (part of) the function of E to the extent that E's having F now supports the fact that tokens of E will continue to exist from now on into the near future.⁷ This can be

⁷ While the focus on tokens of E continuing to exist from now on into the near future is future-focused, the fact that the source of this continued existence is in the present makes it helpful to call this account “presentist.” (This is also

expressed well by borrowing concepts from evolutionary biology (which has a long history in economics and the social science in general—see e.g. Veblen, 1898, Schumpeter, 1959, Hayek, 1973, Nelson and Winter, 2002, Hodgson and Knudsen, 2010, Schulz, 2020, Langlois, 1985, Vromen, 2018, Witt, 2016):

(Presentist Social Functionalism) For any feature F of social entity E that makes it more likely that E will survive or reproduce in the current socio-cultural environment, F is (part of) the function of E.⁸

Put differently: the function of a social entity lies in those of its features that increase the expected reproductive or persistence success in the current socio-cultural environment of the social entity in question.⁹

Note that PSF is a dynamic, non-historical account. On this picture, the fact that social entity E has function F cannot be taken as an explanation of why the collective is in existence now. This makes for a major contrast to the historical accounts of function (Millikan, 1984, Papineau, 1987), according to which E's having function F is due to E having been selected for F in the past. Rather, on PSF, E's having function F is forward looking from the present: it expresses why

in line with the rest of the literature: Nanay, 2014, Schulz, 2025.) At any rate, this is a purely verbal issue—the labeling of the account does not affect its substantive points.

⁸ Note that this account groups together genuine selection—i.e. the heritable differential reproduction of social institutions—and mere sorting—i.e. the differential growth or persistence of social institutions. This is noteworthy, for while it is traditionally thought that genuine selection requires reproduction, recent work on evolutionary theory suggests that there are evolutionary biological reasons for seeing differential persistence and reproduction as merely different instantiations of the same kind of phenomenon (Doolittle, 2014, Price, 1995, Fresco et al., 2018, Hodgson and Knudsen, 2010). This is especially compelling in the context of the social science: while it is not generally plausible to see social institutions as reproducing, the differential *growth or survivorship* of different kinds of social institutions *is* generally plausible.

⁹ In a slightly different form, the focus on which traits are *adaptive* (fitness-enhancing now)—rather than *adaptations* (fitness-enhancing in the past)—to ground function has been put to use in other contexts, too (see e.g. Nanay, 2014); however, the present form of this idea is unique and novel.

we should become more confident that tokens of E will continue to exist from now on. Of course, this does not mean that there is no explanation for why E has a given feature set S now; it is just that this explanation need not lie in its current function (though it may lie in its past function—which may or may not be equal to its current function). This is helpful, as it avoids one of the classic problems of social functionalism: that the appropriate history is often lacking (Pettit, 1996, Schulz, 2025, Elster, 1982).

On the flipside of this is the fact that the function of a social entity changes dynamically with its environment: the fact that E has function F today—since F helps E survive or reproduce today—does not mean that E will also have function F tomorrow. After all, it may be that, due to changes in the environment, it will be feature G that enables E to survive and reproduce then. Indeed, it may be precisely the presence of F that brings about these changes in the environment: a successful fire department may be thus that it leads to the absence of the need to fight fires (see also Satz, 2013).¹⁰ However, this is as it should be: in a world without fires, putting them out should not be seen to be the function of fire department—though preventing them from being started might be. The point to note is that a functional feature of an institution is not what guarantees its long-term survival; a functional feature is what makes it more likely that the institution will survive or reproduce in the here and now. In the famous Keynesian saying: in the long run, we are all dead.¹¹

This also matters here, as it helps to address the second point noted above: concerning the prevalence of functional NISE's. On the one hand, given that social purpose is tied to the

¹⁰ The paradox of democracy involves a similar kind of self-undermining: if a populace votes to abandon voting, then that brings it about that a genuine, voting-based democracy can no longer flourish (or at least not due to its being responsive to the voters' wishes). I thank a referee for helpful discussion of this point.

¹¹ That said, in line with all accounts of selection, a minimal degree of environmental stability is assumed here. If environments change extremely quickly, evolution will effectively follow a Brownian motion process. However, this assumption is innocuous, as most actual societies display this minimal degree of stability.

features that *now* increase the expected survival or reproductive success of a social entity, many more (non-individualist) social entities can be expected to have purposes than may at first seem to be the case. In particular, even newly created social entities can have purposes, and even social entities whose past persistence or reproductive success was heavily driven by chance can still be purposeful now. On the other hand, though, the account is also not trivial or pan-functional: not every social entity always has a purpose. (This will become important again momentarily.)¹²

A further noteworthy feature of PSF especially in the present context is the fact that, according to PSF, a social institution will generally have many functional features of different importance—as well as many disfunctional and neutral ones. It is not that social entity E is restricted to having just one functional feature F. Rather, PSF can allow for F_1 to F_n to all be *part of* the function of E—perhaps with some F_i being particularly important, as they have the biggest single impacts on the expected reproductive or survival success of E. Similarly, the fact that E has F_1 to F_n does not mean that it does not also have disfunctional features G_1 to G_n that decrease its expected reproductive or survival success. In this way, PSF allows for far more fine-grained functional ascriptions to social entities than many other versions of social functionalism.¹³

¹² Relatedly, PSF also spells out the purpose of social entities in a way that does not presuppose that this purpose aims at some human good or specific set of objectively moral ends. This is beneficial, as justifying what these ends are is very difficult. In this way, it PSF differs strongly from Miller's (2017) account of social functionalism, which *is* thickly moral in this sense. Of course, PSF is normative in the sense that all functional ascription is: it describes what a social institution ought to do, and not just what it, in fact, does do (Millikan, 1984, Millikan, 1990, Millikan, 1989, Fodor, 1990). Also, this point is independent of whether functional kinds are real (Craver, 2009, Dupre, 1993, Dupre, 2024, Frasser and Guzmán, 2020).

¹³ We could create a summary measure M of E's function by averaging its functional features, weighted by their relative importance on E's expected reproductive or survival success: $M = \sum_{i=1}^n a_i \times F_i$, where $a_i = \frac{\partial w}{\partial F_i}$, and w is E's expected reproductive or survival success. However, in the present context, there is no need to rely on this measure, so it will not be considered further here. That said, this point is in line with the fact institutional kinds typically exhibit clusters of correlated properties (Boyd, 1999), and are thus often multi-functional kinds (Guala, 2023). It is also noteworthy that this account is not trivial, in that an institution can have many functional features, and can still fail to persist—just like a driver with the fastest F1 car can still lost the race—e.g. because they have a tire puncture or there is a driver mistake.

With PSF in the background, we can go on to characterize the *interest* of a (non-individualist) social entity as follows:

(PSF-Interest) The interest of social entity E is (a) the preservation of those features F_i of E that make it more likely that E will survive or reproduce in the current socio-cultural environment, or (b) the abandonment of those features G_i of E that make it less likely that E will survive or reproduce in the current socio-cultural environment.¹⁴

In other words, the key idea behind this account is that the interest of a social entity lies in safeguarding and enhancing whatever features it has that allow it to survive or reproduce in the here and now. It is in E's interest to make sure that it can function well in its current social environment.

In turn, *PSF-Interest* allows for a genuinely non-individualistic account of the interest of a NISE. The features F that increase the expected reproductive or survival success of E need not be reducible to or derivable from features of individual members of E. They can be genuine features of E: how easily it spreads in the culture, say, how resilient it is to change, or how deeply embedded it is in the society in question.¹⁵ Again, it is crucial to emphasize that, as such, other theories of social functionalism could also be used to assign *interests* to social entities. Still, relying on PSF to ground the interests of NISE has two key benefits (apart from the general ones attached to PSF just mentioned).

¹⁴ In line with the discussion in note 13, we could see an institution as harboring competing interests in the sense that, in different contexts, the preservation of different functional features may be emphasized. However, this is only a mild internal conflict, in that the institution's overall interest remains in preserving *all* of its functional features (and suppressing all of the disfunctional ones). See also section IV.

¹⁵ These features can also be independently measurable, so that we do not need to rely on past data about which institutions survived or reproduced to assign functional features to them. This thus further underwrites the fact that the account is non-trivial (see also note 13).

First, by relying on PSF, we are able to accommodate the individualistic approach to the interests of NISE *where the latter is plausible*. According to the account defended here, there may be nothing problematic about seeing the interest of a collective *sometimes* in the interest of the majority of its members. After all, it may often be precisely this fact that increases the expected reproductive or survival success of that institution: sports clubs, say, that are not sensitive to the views of the majority are in danger of ceasing to have members, and thus dying out. The same goes for power relationships—i.e. *who* the institution is for (Emerson, 1962, Emerson, 1976, Weber, 1968). Favoring certain individuals over others can, at times, increase an institution's expected reproductive or survival success (e.g. if these individuals have significant influence in the society in question). Similarly, it is sometimes thought that the purpose of a corporation—and thus its interests—lie in maximizing profits or shareholder value (see e.g. Friedman, 1970). Whatever else may be said about these ideas (see e.g. Friedman, 1970, Freeman, 1984, Orts and Strudler, 2002), the important point to note here is just that the *PSF-Interest*-based account can accept this conclusion—as long as it is case that increasing profits or shareholder value increases the expected survival or reproductive success of corporations. (Importantly, where this is not the case, it seems clear that it *PSF-Interest* that gives the right answer about the purpose of the corporation: it just does not seem relevant whether a corporation does or does not maximize profits if this is unimportant for its expected reproductive or persistence success.) In this way, the account here laid out can accept the best of the individualistic account, all the while avoiding its downsides.

However, second, the present account can go beyond the individualist account as it is not *committed* to the interests of a NISE being reducible to those of its (current) members. In particular (as noted earlier), *PSF-Interest* can be applied to collectives and social institutions—

such as those of non-profits and democracies—that cannot be analyzed in terms of the maximization of profits or shareholder values (as they lack these features). In this way, *PSF-Interest* can be used to make sense of the kind of COI the previous accounts have overlooked. In particular, we can now see that there can be cases of COI that do not involve two individual agents, but where the conflict is between an individual and a NISE—such as a firm or other social institution. This occurs when the interests of the individual and those of the NISE are not fully aligned. As noted earlier, there is good reason to think that these kinds of NICOI are very widespread. In particular, there will, in general, be many cases where an individual’s interests do not (always) match what would aid the relevant NISE to increase *its* expected reproductive or survival success. Importantly, also this theory can now be made mathematically precise. Indeed, we can use the same framework familiar from standard formulations of the PAP to investigate this.

So, consider a corporation C and its (sole current) employee / owner M. Assume further that it is in C’s interest to make high-quality products that last a long time (H): this leads customers to stick with the company the next time they are in the relevant market, which in turn increases the medium- to long-term longevity of C. Assume further, though, that it is in M’s interest to make cheaper products that break more often: M’s time-horizon is shorter than that of C—M will not be an employee / owner when customers are likely to be back in the relevant market—and building high-quality product is more costly for M. The corporation can have costly but robust quality controls (R), or a cheaper, but more cursory such controls (\sim R). In that case, a COI can arise, as the interests of the manager and the corporation as a whole are not fully aligned. Figure 2 spells out a simple model of this kind of conflict:

Employee / Corporation	R	~R
H	1, 2	1, 3
~H	0, -1	2, -2

[Figure 2: a corporate institutional conflict of interest]

The key thing to note here is that we can now recognize this situation to have the same structure as the work / shirk game from earlier. This is explanatorily important, as there is only one individual person here (the employee / owner)—but it now becomes clear that there could still be a COI here. Similar to what was the case earlier, the equilibrium here involves “inspecting” (quality control) with probability $p = \frac{1}{2}$ and the employee switching between working and shirking with probability $q = \frac{1}{2}$.

The one difference is that the “choice” of the corporation as to what kind of quality control to employ needs to be interpreted slightly differently. The corporation faces the choice of which quality control protocol to adopt: it can always check ($q = 1$), never check ($q = 0$), or anything in between. Importantly, while these quality control protocols *may* be adopted or carried out by a human actor, they need not be: they might be part of the manufacturing machinery or carried out by an AI. The key point is just that these protocols are features of the corporation: these are different ways for the corporation to be (however, exactly, that is implemented). The corporation’s choice consists in the fact that it needs to assess which one to adopt. Importantly, this assessment can be made by a cultural evolutionary process: different quality protocols are differentially helpful for the corporation to survive or reproduce.¹⁶ Of course, human decisions can be involved in determining whether the corporation sticks with a particular quality control

¹⁶ This fits the fact that many corporations are known to go through constant cycles of “reworks”—from cost-cutting to expansion and back again; see also Witt and Worch (2023).

protocol—but these human decisions are themselves subject to the cultural evolutionary assessment. That is to say, it is the corporation’s interests, as set out by *PSF-Interest*, that determines which quality control schemes it is best to adopt.

Without a doubt, it needs to be acknowledged that this way of understanding choice is different from the choice a human employee might make.¹⁷ However, for two reasons, this is not problematic. On the one hand, this sense of choice is not unprecedented: indeed, it is familiar from work in evolutionary game theory, evolutionary ecology, or population genetics.¹⁸ Take an organism like a penstemon flower; it faces the choice of which color its pedals should have. Assuming it is in an area frequented by hummingbirds, a choice of red or orange is being rewarded, as this “pops out” to the hummingbirds (its pollinators) (Wessinger et al., 2023). However, of course, this “choice” amounts to the fact that plants with the right genes are being favored by selection (though depending on the case, it need not be genetically driven—as is e.g. true when it comes to mycorrhizal symbioses). Of course, penstemon flowers do not choose pedal color like a human might choose the color of their shirt. This is a choice that is being made and assessed by an evolutionary process. Much the same is going in the case of corporation.

Importantly and on the other hand, while this seems quite different from an intentional, human choice, it is not an unreasonable sense of choice. In the first place, it still involves the evaluation of an option space. Second, many human choices are also habitual and subconscious (Kahneman et al., 1982). Third and most importantly, this difference in the nature of choice is in line with the fact that we are trying to model a different kind of phenomenon: NICOI. Put differently, the key is that with *PSF-Interest*, we can model situations using the principal / agent

¹⁷ I thank an anonymous referee for very helpful discussion of this point.

¹⁸ Falster and Westoby (2003) write: “Biological strategies are construed as phenotypes [...]. Situations are called games when the viability of a strategy depends upon what strategies are being ‘played’ by competitors.”

framework that could not be modeled in this way before: we can now recognize that these otherwise different situations amount to a game with the same structure. Of course, this is an *extension* of work on PAPs—but it is precisely the fact that this *is* an extension that makes it the case that the explanatory payoff of PAPs is being increased considerably here.¹⁹

In short: by using *PSF-Interest* to spell out the interests of NISE like firms and other social institutions, NICOI can be modeled and theoretically handled. This then naturally leads to the question of *when* it is reasonable to treat a collective as a player of its own—i.e. how to diagnose whether a NICOI may be occurring in a given situation.²⁰ Now, on an abstract level, the framework here laid out gives a straightforward answer: a collective should be seen to have interests of its own—and thus to be a potential party to a conflict of interest—when it faces its own, separate cultural evolutionary pressures in the present environment. However, there are also some concrete diagnostics that can be derived from the above discussion, and which can be used as a framework with which to infer when a collective should be modeled as an actor, rather than merely a part of the “incentive landscapes” of individual agents. Four key such diagnostics for treating a collective as a NISE and thus (potentially) party to a NICOI are the following:²¹

¹⁹ The present account could also be seen in line with classic insider / outsider dynamics—unions, etc. (Lindbeck and Snower, 1988, Lindbeck and Snower, 2001). In the case of unions, there can be conflict between “insiders” (union members) who want to ensure a higher wage by restricting labor supply and “outsiders” (other workers), who may have been willing to work at a lower wage rate. It is just that, in the present case, the “outsiders” are merely possible future people. Since, as noted earlier, such merely possible people do not have interests, the account in the text based on NISE’s is superior. (The same goes for the paradox of democracy mentioned earlier). (Of course, it may be that the insiders anticipate the needs of the outsiders, and act accordingly—but this is not guaranteed. This point was anticipated by Knight, 1921, see also Emmett, 2021).

²⁰ This issue is related to the discussion of the major transitions and MLS in evolutionary biology—with similar upshots (Maynard Smith and Szethmary, 1995, Okasha, 2006, Sober and Wilson, 1998, Godfrey-Smith, 2009).

²¹ These diagnostics will often be correlated. However, it is possible that they come apart, in which case the situation would need to be assessed by closely analyzing its specific details.

- *Time Horizon*: The time horizon of the collective is different from that of the individual (individual managers may be evaluated / rewarded on a yearly basis, whereas corporations may only be evaluated / rewarded over decades).
- *Composition*: The collective comprises non-actually existing individuals (US democracy extends beyond the current voters—as opposed to Mr. Smith’s 5th grade class, which only includes the students in the class this year).
- *Functional Specialization*: There is a lot of functional specialization and division of labor in the collective (the features of the collective are more than the sum of their parts, but depend on their integration as well).²²
- *Behavioral Routines & Rules*: The collective operates by means of culturally inherited behavioral routines / institutions (and thus, the relevant actions are not just features of the individuals in question).

Before considering some further benefits of this framework, it is helpful to be clear about the structure of the argument here. This argument has two components. First, situations like the above of the corporation and the owner / employee *can* be seen as involving a genuine NICOI. While the notion of institutional choice is different, it still sufficiently well-grounded to make the extension of the principal / agent framework possible here. Importantly also, the paper shows that this can all be done in a “naturalized” manner: institutional interests are not some mysterious entity, but grounded in a well-established cultural evolutionary process. Second, it is not just *possible*, but in fact *plausible*, to see these kinds of situations as cases of NICOI. While it may (often) also be possible to see these cases as just involving individual agents that make decisions

²² See Hoover (2010) on the distinction between “natural” and “synthetic” aggregates.

using evolved rules of thumb that have unintended consequences (see also notes 3 and 5 for more on this), treating these as genuine NICOI allows researchers to recognize that these phenomena share key features with classic PAPs. The next section brings out some of these benefits in more detail.²³

IV. An Application: An Expanded Taxonomy of Internal Conflicts in Collective Agents

This new *PSF-Interest*-based account of NICOI can make more precise how to think about—and manage—some of the problems that can arise as people cooperate and coordinate their actions. Several different applications come to mind here, from cases of institutional corruption (Thompson, 1995, Lessig, 2013, Miller, 2017, Schulz, 2023) to the legal framework surrounding COI (Fama and Jensen, 1983). To give the discussion focus, though, the rest of this section will concentrate on *internal conflicts in collective agents*.

In general, the problem of internal conflicts in collective agents concerns the fact that the conflict between the need for cooperation and our egoistic tendencies should not be seen to be a one-shot affair. Rather, we often find cooperation and coordination and our egoistic tendencies to be coexisting: even where we, as individual agents, manage to cooperate and coordinate to achieve collective actions, internal conflicts can persist. In line with the individualistic focus of the work on COI, though, the study of the nature of internal conflicts in collective agents (and the ways of managing them) has tended to focus on conflicts between different individuals internal to a given collective agent. However, given the above framework for the study of NICOI, this can now be seen to be overly limited.

²³ A related set of issues arise in evolutionary biology more generally. On purely gene-centric views, we should analyze all evolutionary phenomena from what benefits different sets of genes. By contrast, from a multi-level selectionist perspective, we can recognize evolutionary entities on many levels of the biological hierarchy. See also note 20.

To see this, it is best to begin by clarifying the nature of the problem. For concreteness, it is useful to focus on the case of a *firm* (though parallel remarks could be made for other economic collectives, such as government agencies, NGO's, not-for-profits, and the like). A typical firm has a number of employees; these employees are likely to have at least somewhat differing interests from each other. In turn, this gives rise to the possible existence of conflicts within the firm: it is not just the case that the firm may have COI with other economic actors—there may be COI within the firm itself.²⁴

The most obvious such case concerns instances of the work / shirk game of figure 1. It is presumed that both the manager and the employee recognize that it is in their interest to play this “game” (of being part of the same corporation), but they differ over what the best outcome is of this game. In this case, then, the internal conflict in the firm is centered on the fact that different individuals inside the firm face COI. This, though, is not the only kind of internal conflict widely studied in economics.

The other major such internal conflict centers on public goods games (see e.g. Falk and Fischbacher, 2006, Reeson and Tisdell, 2008, Fehr and Schmidt, 1999, Fehr and Gaechter, 2000, Henrich et al., 2001, Falk et al., 2003, Fehr and Fischbacher, 2003). In the work / shirk-type of internal conflicts, the different elements of the collective structure in question (the firm, say) are directly playing against each other. While not necessarily zero-sum games, these are still situations where one's individual's gain is another individual's loss. It is inherent to these games that the principal and the agent are playing against each other: the principal wants the agent to do things that the agent does not want to do (at least in this way). It is the last point that is relaxed in

²⁴ What follows merely presumes that firms are collective agents, in the sense of being integrated collectives of individual economic agents. Beyond this, no further assumptions about the nature of collective agency or firms are made (Nelson and Winter, 1982, Nelson et al., 2018, Knight, 1921, Hodgson, 1999, Coase, 1937, Coase, 1960, Winter, 1988, Hart and Holmström, 1987).

the other major existing framework used to explore internal conflicts. In the latter, public good-focused kind of cases, it is not true that one person's gaining an advantage implies that someone else loses. On the contrary, by contributing to the public good, we can all benefit. The conflict here consists in the fact that different individuals may contribute different amounts. In particular, some agents may choose to free-ride on the contributions of others.²⁵ That is, they may decide to consume the public good without adding to it. For example, each subject may receive a set number of monetary units that they can invest into a common pool. The total contributions in the pool are then multiplied by fixed amount n and split evenly among the subjects. In public goods games, the multiplication constant n of the common pool ensures that each subject would prefer it if everyone else contributed to the common pool—except for themselves. Depending on the details of the case—whether people can interact multiple times, punish each other, etc.—more or less contribution to the public good will result.

For present purposes, it is key to note that while these public goods-based cases do not feature individuals with interests that are directly opposed to each other, we are here still maintaining the focus on inter-individual interactions that was true of the PAPs. In both cases, the source of the conflict is between the different agents making up the collective in question. Given the framework of NICOI developed in the previous section, this can now be seen to overly limited, though.

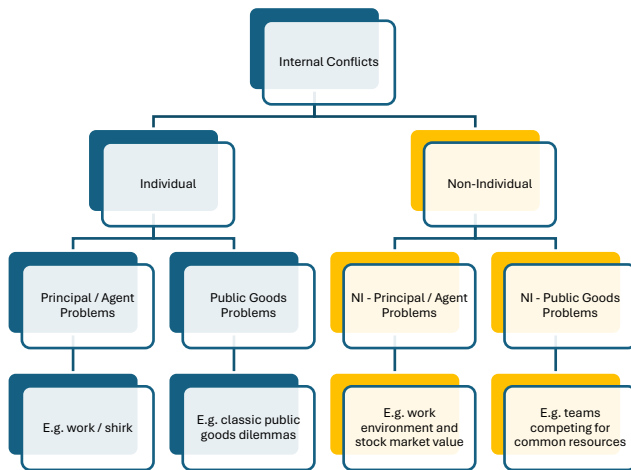
In particular, there are two other kinds of internal conflicts that can now be seen more clearly. On the one hand, there are internal conflicts between an employee (or other individual member) of the firm (or other collective in question) and the firm itself. This is the kind of case depicted in

²⁵ More complex, multi-dimensional versions of this case could be considered as well, in which agents may prefer to use the common resources for one purpose, others for another (or favor certain resources over others). For present purposes, though, considering these further complexities is not so central, though.

figure 2 above: what the firm “wants” may conflict with what its particular individual members might want.

Importantly and on the other hand, this kind of COI could even occur just among NISE’s. For example, we could have a larger firm with several teams. The interests of one team—i.e. what, in the context of the firm environment, might help that team expand or survive—may not be what is in the interest of the firm as a whole. For example, it may be in the interest of a specific team to centralize as many activities under its auspices as possible—even though that may lead to a less efficient division of labor than if these activities were spread more widely among different units. Something like this could also lead to non-individualistic versions of a public goods game: it might be that all the teams would profit if they shared common pool resources efficiently, but each team individually prefers to free-ride on the public contributions of other teams.²⁶

Putting all of this together, this thus yields an expanded taxonomy of internal conflicts:



[Figure 3: a taxonomy of internal conflicts; the orange parts are new additions]

²⁶ Indeed, this point is familiar from some classic work in political economy, where state actors (the US, the Soviet Union, etc.) are treated as agents of their own—e.g. when considering the “Bay of Pigs” invasion. The point in the text is that this can now be seen as an instance of typical COI—and thus modeled in the same way. I thank an anonymous referee for much useful discussion of this point.

In this way, it becomes clear that it is possible to capture a wider range of internal conflicts using familiar tools—i.e. those of the principal / agent framework and public goods games. Put differently, the account laid out here allows us to recognize and predict the existence of internal conflicts that we would otherwise overlook.²⁷ This thus brings out some of the benefits of the perspective advocated here: moving beyond seeing institutions merely as structures that constrain individual actions, but as *parties* to a conflict of interest can make it easier to see how to model and analyze these cases.

Importantly also, with this taxonomy in the background, it becomes possible to make some clear predictions about the kinds of cases in which non-individualist internal conflicts are particularly severe—and how to prevent that from happening. This is a further benefit of the perspective advocated here.

It is a basic insight of the principal / agent framework that the more (fewer) control mechanisms there are, the less (more) pronounced the PAP will be. Similarly, the more ways there are of forcing contribution to public goods (e.g. through punishment), the less such internal conflicts there are. This matters, as the availability of such processes can now be recognized to be relevant to a wider set of phenomena—as a wider set of phenomena can now be recognized to be COI's.

So, as noted in the discussion of figure 2, there may be different versions of a corporation C—or there may be different “behavioral routines” concerning compensation in C—so that M's compensation is more or less heavily skewed towards stock grants. This can then set up a

²⁷ Another milder form of internal conflict occurs when different teams or individuals within an institution (a corporation, say) favor enhancing or preserving different functional aspects of the institution. However, this is only a minor internal conflict, in that all of these different viewpoints still lead to an increase in the institution's expected reproductive or survival success—the issue only concerns the *strength* of this increase. See also note 14.

cultural evolutionary process (Henrich, 2000, Boyd and Richerson, 2005, Nelson and Winter, 1982, Nelson and Winter, 2002) that selects for compensation schemes that favor the interests of the corporation as a whole. For example, the employee's compensation scheme could be based in part on the long-term quality of the product. It is of course also possible for the corporation to fail to mitigate its internal conflicts properly, and collapse.

This kind of case is not restricted to the economic sphere, and can also occur in the political realm. For example, consider the case of governments—whether democratically elected or autocratic—making decisions that are not in the interest of the state. So, a leader may make decisions that are in their interests (which may include enriching themselves, but could also just be centered on advancing their personal moral ideology), even though these decisions do not in fact help the state as a whole thrive. So, perhaps what would benefit the state is a strong commitment to the rule of law and due process, whereas the leader prefers to make decisions in line with their convictions, independently of the procedural requirements. Importantly, these decisions need not be opposed or go against the interests of anyone in the society—however, they are still in opposition to the interest of the state as a whole though. As noted earlier in the context of the “paradox of democracy,” “the people” can be wrong about what is good for the state as a whole. Importantly, as was the case in the economic sphere, different mitigation strategies could evolve that the state can “use” to ameliorate the situation here. These include constitutional mechanisms that enshrine due process into the fabric of political decision-making, as well as robust systems of checks and balances.

V. Conclusion

This paper has shown that, by appeal to a contemporary theory of social functionalism—such as Presentist Social Functionalism—non-individualistic forms of conflicts of interest can be grasped with the standardly available tools—including, in particular, the principal / agent framework. More specifically, the paper has shown that the interests of a non-individualistic social entity—such as a social institution or collective—can be seen to rest in what allows it to increase its expected reproductive or survival success in the current socio-cultural-economic environment. This makes clear that the two main types of internal conflicts —classic PAPs and public goods dilemmas—come in *two* versions each: an individualistic and a non-individualistic one. The last case has not been as clearly recognized before now.

Bibliography

- AOKI, M. 2001. *Toward a Comparative Institutional Analysis*, Cambridge, MA, MIT Press.
- ARROW, K. 1951. *Social Choice and Individual Values*, New York, Wiley.
- BOYD, R. 1999. Kinds as the ‘Workmanship of Men’: Realism, Constructivism, and Natural Kinds. In: NIDA-RUMELIN, J. (ed.) *Rationalitat, Realismus, Revision: Proceedings of the Third International Congress, Gesellschaft fur Analytische Philosophie*. Berlin: de Gruyter.
- BOYD, R. & RICHERSON, P. 2005. *The Origin and Evolution of Cultures*, Oxford, Oxford University Press.
- BUCHEN, C. 2024. Institutional resilience: how the formal legal system sustains informal cooperation. *Journal of Institutional Economics*, 20, 1–15.
- COASE, R. 1937. The Nature of the Firm. *Economica*, 4, 386-405.
- COASE, R. 1960. The problem of social cost. *Journal of Law and Economics*, 3, 1-44.
- CRAVER, C. 2009. Mechanisms and natural kinds. *Philosophical Psychology*, 22, 575-594.
- DEMSKI, J. S. 2003. Corporate Conflicts of Interest. *Journal of Economic Perspectives*, 17, 51–72.
- DOOLITTLE, W. F. 2014. Natural selection through survival alone, and the possibility of Gaia. *Biology and Philosophy*, 29, 415-423.
- DUPRE, J. 1993. *The disorder of things: Metaphysical foundations of the disunity of science*, Cambridge, MA, Harvard University Press.
- DUPRE, J. 2024. The disunity of science and the unity of the world. *Philosophy of Science*, 91, 1043-1057.
- ELSTER, J. 1982. The Case for Methodological Individualism. *Theory and Society*, 11, 453-482.

- EMERSON, R. M. 1962. Power-dependence relations. *American Sociological Review*, 27, 31-41.
- EMERSON, R. M. 1976. Social exchange theory. *Annual Review of Sociology*, 2, 335-362.
- EMMETT, R. B. 2021. Uncertainty and the social organization of economic activity. *Journal of Institutional Economics*, 17, 883-895.
- EPSTEIN, B. 2015. *The Ant Trap: Rebuilding the Foundations of the Social Sciences*, Oxford, Oxford University Press.
- FALK, A., FEHR, E. & FISCHBACHER, U. 2003. On the Nature of Fair Behavior. *Economic Inquiry*, 41, 20-26.
- FALK, A. & FISCHBACHER, U. 2006. A theory of reciprocity. *Games and Economic Behavior*, 54, 293-315.
- FALSTER, D. S. & WESTOBY, M. 2003. Plant height and evolutionary games. *Trends in Ecology & Evolution*, 18, 337-343.
- FAMA, E. & JENSEN, M. 1983. Agency Problems and Residual Claims. *Journal of Law & Economics*, 26, 327-350.
- FEHR, E. & FISCHBACHER, U. 2003. The nature of human altruism. *Nature*, 425, 785-791.
- FEHR, E. & GAECHTER, S. 2000. Fairness and Retaliation: The Economics of Reciprocity. *The Journal of Economic Perspectives*, 14, 159-181.
- FEHR, E. & SCHMIDT, K. M. 1999. A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, 114, 818-868.
- FODOR, J. 1990. *The Theory of Content*, Cambridge, MA, MIT Press.
- FRASSER, C. & GUZMÁN, G. 2020. What do we call money? An appraisal of the money or non-money view. *Journal of Institutional Economics*, 16, 25-40.
- FREEMAN, R. E. 1984. *Strategic Management: A Stakeholder Approach*, Boston, Pitman.

- FRESCO, N., JABLONKA, E. & GINSBURG, S. 2018. The Construction of Learned Information Through Selection Processes. *In: JOYCE, R. (ed.) The Routledge Handbook of Evolution and Philosophy*. New York: Routledge.
- FRIEDMAN, M. 1970. The Social Responsibility of Business is to Increase Its Profits. *The New York Times Magazine*.
- GLIMCHER, P. W., DORRIS, M. C. & BAYER, H. M. 2005. Physiological utility theory and the neuroeconomics of choice. *Games Econ Behav*, 52, 213-256.
- GODFREY-SMITH, P. 2009. *Darwinian Populations and Natural Selection*, Oxford, Oxford University Press.
- GREIF, A. & KINGSTON, C. 2011. Institutions: Rules or Equilibria? *In: SCHOFIELD, N. & CABALLERO, G. (eds.) Political Economy of Institutions, Democracy and Voting*. Berlin: Springer.
- GROSSMAN, S. & HART, O. 1983. An Analysis of the Principal-Agent Problem. *Econometrica*, 51, 7-45.
- GUALA, F. 2016. *Understanding Institutions: The Science and Philosophy of Living Together*, Princeton, Princeton University Press.
- GUALA, F. 2023. Social kinds: historical and multi-functional. *European Journal for Philosophy of Science*, 13, 1-15.
- HART, O. & HOLMSTRÖM, B. 1987. The theory of contracts. *In: BEWLEY, T. (ed.) Advances in Economics and Econometrics*. Cambridge: Cambridge University Press.
- HAYEK, F. 1973. *Laws Legislation and Liberty, Vol. 1: Rules and Order*, London, Routledge.

- HENRICH, J. 2000. Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining Among the Machiguenga of the Peruvian Amazon. *American Economic Review*, 90, 973-979.
- HENRICH, J., BOYD, R., BOWLES, S., CAMERER, C., FEHR, E., GINTIS, H. & MCELREATH, R. 2001. In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *The American Economic Review*, 91, 73-78.
- HODGSON, G. 1999. *Evolution and Institutions*, Cheltenham, Edward Elgar.
- HODGSON, G. 2007. Meanings of Methodological Individualism. *Journal of Economic Methodology*, 14, 211-226.
- HODGSON, G. & KNUDSEN, T. 2010. *Darwin's Conjecture*, Chicago, University of Chicago Press.
- HOOVER, K. D. 2009. Microfoundations and the Ontology of Macroeconomics *In: ROSS, D. & KINCAID, H. (eds.) The Oxford Handbook of Philosophy of Economics*. Oxford: Oxford University Press.
- HOOVER, K. D. 2010. Idealizing Reduction: The Microfoundations of Macroeconomics. *Erkenntnis*, 73, 329-347.
- JONES, T. 1996. Methodological Individualism in Proper Perspective. *Behavior and Philosophy*, 24, 119-128.
- KAHNEMAN, D., SLOVIC, P. & TVERSKY, A. 1982. *Judgment under Uncertainty: Heuristics and Biases*, Cambridge, Cambridge University Press.
- KNIGHT, F. 1921. *Risk, Uncertainty, and Profit*, Boston, MA, Houghton Mifflin.
- LANGLOIS, R. 1985. Rationality, Institutions, and Explanation. *In: LANGLOIS, R. (ed.) Economics as a Process*. Cambridge: Cambridge University Press.

- LESSIG, L. 2013. "Institutional corruption" Defined. *Journal of Law, Medicine & Ethics*, 413, 553-555.
- LINDBECK, A. & SNOWER, D. J. 1988. *The Insider-Outsider Theory of Employment and Unemployment*, Cambridge, MA, MIT Press.
- LINDBECK, A. & SNOWER, D. J. 2001. Insiders versus Outsiders. *Journal of Economic Perspectives*, 15, 165-188.
- LIST, C. 2012. The theory of judgment aggregation: An introductory review. *Synthese*, 187, 179-207.
- LIST, C. & PETTIT, P. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*, Oxford, Oxford University Press.
- LITTLE, D. 1998. *Microfoundations, Method, and Causation: On the Philosophy of the Social Sciences*, New Brunswick, NJ, Transaction Publishers.
- MAS-COLELL, A., WHINSTON, M. D. & GREEN, J. R. 1995. *Microeconomic Theory*, Oxford, Oxford University Press.
- MAYNARD SMITH, J. & SZETHMARY, E. 1995. *The Major Transitions in Evolution*, Oxford, Oxford University Press.
- MERTON, R. 1968. *Social Theory and Social Structure*, New York, Free Press.
- MILLER, S. 2017. *Institutional Corruption: A Study in Applied Philosophy*, Cambridge, Cambridge University Press.
- MILLIKAN, R. 1984. *Language, Thought, and Other Biological Categories*.
- MILLIKAN, R. 1989. Biosemantics. *Journal of Philosophy*, 86, 281-297.
- MILLIKAN, R. 1990. Truth Rules, Hoverflies, and the Kripke-Wittgenstein Paradox. *The Philosophical Review*, 99, 323-353.

- NANAY, B. 2014. Teleosemantics without Etiology. *Philosophy of Science*, 81, 798-810.
- NELSON, R., DOSI, G., HELFAT, C. E., PYKA, A., SAVIOTTI, P. P., LEE, K., DOPFER, K.,
MALERBA, F. & WINTER, S. (eds.) 2018. *Modern Evolutionary Economics: An
Overview*, Cambridge: Cambridge University Press.
- NELSON, R. & WINTER, S. 1982. *An Evolutionary Theory of Economic Change*, Cambridge,
MA, Belknap Press.
- NELSON, R. & WINTER, S. 2002. Evolutionary Theorizing in Economics. *Journal of
Economic Perspectives*, 16, 23-46.
- OKASHA, S. 2006. *Evolution and the Levels of Selection*, Oxford, Oxford University Press.
- ORTS, E. W. & STRUDLER, A. 2002. The Ethical and Environmental Limits of Stakeholder
Theory. *Business Ethics Quarterly*, 12, 215-233.
- PAPINEAU, D. 1987. *Reality and Representation*, Oxford, Blackwell.
- PARSONS, T. 1951. *The Social System*, London, Routledge.
- PETTIT, P. 1996. Functional Explanation and Virtual Selection. *The British Journal for the
Philosophy of Science*, 47, 291-302.
- POPPER, K. 1945. *The Open Society and Its Enemies*, London, Routledge.
- PRICE, G. R. 1995. The Nature of Selection. *Journal of Theoretical Biology*, 175, 389-396.
- REESON, A. F. & TISDELL, J. G. 2008. Institutions, motivations and public goods: An
experimental test of motivational crowding. *Journal of Economic Behavior &
Organization*, 68, 273-281.
- ROSS, D. 2005. *Economic Theory and Cognitive Science: Microfoundations*, Cambridge, MA,
MIT Press.

- ROSS, S. A. 1973. The Economic Theory of Agency: The Principal's Problem. *The American Economic Review*, 63, 134-139.
- RUBEN, D.-H. 1985. *The Metaphysics of the Social World*, London, Routledge.
- RUIZ, N. & SCHULZ, A. 2023. Microfoundations and Methodology: A Complexity-Based Reconceptualization of the Debate. *British Journal for the Philosophy of Science*, 74, 359-379.
- SATZ, D. 2013. Markets, Privatization and Corruption. *Social Research*, Vol. 80, 993-1008.
- SCHOTTER, A. 1981. *The Economic Theory of Social Institutions*, Cambridge, Cambridge University Press.
- SCHULZ, A. 2020. *Structure, Evidence, and Heuristic: Evolutionary Biology, Economics, and the Philosophy of their Relationship*, New York, Routledge.
- SCHULZ, A. 2022. What's the Point? A Presentist Social Functionalist Account of Institutional Purpose. *Philosophy of the Social Sciences*, 52, 53-80.
- SCHULZ, A. 2023. Institutional Corruption: The Teleological and Nonnormative Account. *Journal of Ethics and Social Philosophy*, 25.
- SCHULZ, A. 2025. *Presentist Social Functionalism*, Cham, Springer.
- SCHUMPETER, J. 1959. *The Theory of Economic Development*, Cambridge, MA, Harvard University Press.
- SEARLE, J. 1995. *The Social Construction of Reality*, New York, Free Press.
- SEARLE, J. 2005. What is an institution? *Journal of Institutional Economics*, 1, 1-22.
- SEN, A. 1970. *Collective Choice and Social Welfare*, San Francisco, Holden-Day.
- SKYRMS, B. 2004. *The Stag Hunt and the Evolution of Social Structure*, Cambridge, Cambridge University Press.

- SMITH, A. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*.
- SOBER, E. & WILSON, D. S. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge, MA, Harvard University Press.
- THOMPSON, D. F. 1995. *Ethics in Congress: From Individual to Institutional Corruption*, Washington, DC, Brookings Institution.
- VEBLEN, T. 1898. Why is Economics not an Evolutionary Science? *The Quarterly Journal of Economics*, 12, 373-397.
- VROMEN, J. 2018. Generalized Darwinism in Evolutionary Economics: The Devil Is in the Detail. In: WITT, U. & CHAI, A. (eds.) *Understanding Economic Change: Advances in Evolutionary Economics*. Cambridge: Cambridge University Press.
- WATKINS, J. 1952. The Principle of Methodological Individualism. *The British Journal for the Philosophy of Science*, 3, 186-189.
- WEBER, M. 1968. *Economy and Society: An Outline of Interpretive Sociology*, New York, Bedminster Press.
- WESSINGER, C. A., KATZER, A. M., HIME, P. M., RAUSHER, M. D., KELLY, J. K. & HILEMAN, L. C. 2023. A few essential genetic loci distinguish *Penstemon* species with flowers adapted to pollination by bees or hummingbirds. *PLoS Biol*, 21, e3002294.
- WINTER, S. 1988. On Coase, Competence, and the Corporation. *Journal of Law, Economics, and Organization*, 4, 163-180.
- WITT, U. 2016. *Rethinking Economic Evolution – Essays on Economic Change and its Theory*, Cheltenham, Edward Elgar.
- WITT, U. & WORCH, H. 2023. Growth-induced crises and transitions in the governance of firm organizations. *Journal of Economic Behavior & Organization*, 212, 1182-1191.

