

To appear in *Philosophical Explorations*, 14, 2011, pp. 195-217.

Norms and Conventions

Nicholas Southwood and Lina Eriksson

Abstract: What is the relation between norms (in the sense of ‘socially accepted rules’) and conventions? A number of philosophers have suggested that there is some kind of conceptual or constitutive relation between them. Some hold that conventions are or entail special kinds of norms (the ‘conventions-as-norms thesis’). Others hold that at least some norms are or entail special kinds of conventions (the ‘norms-as-conventions thesis’). We argue that both theses are false. Norms and conventions are crucially different conceptually and functionally in ways that make it the case that it is a serious mistake to try to assimilate them. They are crucially different conceptually in that whereas conventions are not normative and are behaviour-dependent and desire-dependent, norms are normative, behaviour-independent and desire-independent. They are crucially different functionally in that, whereas conventions principally serve the function of facilitating coordination, norms principally serve the function of making us accountable to one another.

In ordinary English, the terms ‘norm’¹ and ‘convention’ are sometimes used more or less interchangeably (Gilbert 1989, 2008). Thus, we might speak either of the ‘norm’ or the ‘convention’ among Oxford dons that one must pass the port to the left. Again, we might speak either of the ‘norm’ or the ‘convention’ among the French that one must call strangers ‘vous’ rather than ‘tu’. Talk of ‘conventions’ in these contexts, like talk of ‘norms’, is meant to be explicitly normative. It is meant to refer to particular rules or requirements that are widely accepted within particular social groups: a rule mandating passing the port to the left that is widely accepted among Oxford dons; a rule mandating calling strangers ‘vous’ rather than ‘tu’ that is widely accepted among the French; and so on.

But talk of ‘conventions’ is also sometimes used in a more reductive way, to refer to special kinds of behavioural regularities that are explained by certain

¹ We are only interested here in ‘norms’ in the sense of ‘socially accepted rules or requirements’. See below, section IB.

interlocking patterns of non-normative psychological attitudes. In David Lewis's (1969) classic account, conventions in this sense are solutions to coordination problems.² A coordination problem is a situation in which individuals' interests are aligned such that each desires that everyone behaves in the mutually advantageous way; and coordination is a problem because there are different ways in which they might do that. Examples of coordination problems include the situation of two friends who wish to meet for a drink after work; the situation of the members of a string quartet, each of whom has an interest in everyone playing at the same tempo and in the same key; and a society of drivers, each of whom has an interest in every driver driving on the same side of the road. Conventions are behavioural regularities, sustained by certain interdependent patterns of non-normative beliefs and desires, that constitute solutions to recurring problems of coordination.

We might wonder what is the relation between norms and conventions in this sense? Despite certain *prima facie* differences, a number of philosophers have suggested that there is some kind of significant conceptual or constitutive relation between norms and conventions. Some have argued that conventions are or entail special kinds of norms. Lewis (1969, 99), for example, holds that '[a]ny convention is ... a socially enforced norm: one is expected to conform, and failure to conform tends to evoke unfavourable responses from others'. Let us call this the 'conventions-as-norms thesis'. Others have argued that at least some norms are or entail special kinds of conventions. Let us call this the 'norms-as-conventions thesis'.³ The norms-as-

² It has also been importantly refined by subsequent writers, including Sugden (1986), Marmor (1996), Vanderschraaf (1995), Miller (2001), Millikan (2005), Bicchieri (2006), Verbeek (2002).

³ Adherents of some version of the norms-as-conventions thesis include Edna Ullmann-Margalit (1977); H. Peyton Young (1998, 144-5); Eric A. Posner (1998; 2000); James Coleman (1990); Thomas Voss (2001, 108-9); Bruno Verbeek (2002). Cristina Bicchieri (2006, 26), while not an adherent of the norms-as-conventions thesis, endorses a related idea, namely, that social norms (insofar as they are

conventions thesis might, in principle, be proposed as a way of understanding norms in general, including moral norms (Verbeek 2002).⁴ But it is more commonly (and plausibly) proposed as a thesis about *social norms* in particular.⁵ Thus, the social norm in Australia that people drive on the left is taken to be, or at least to entail, the existence of a drive-on-the-left convention among Australians. The social norm in France that strangers call one another ‘vous’ rather than ‘tu’ is taken to be, or at least to entail, the existence of a vous-rather-than-tu convention among the French. And so on.

We shall argue that both theses are false. Norms (including social norms) and conventions are crucially different conceptually and functionally, such that it is a mistake to try to assimilate them: either by holding that conventions are or entail (special kinds of) norms; or by holding that norms are or entail (special kinds of) conventions. They are crucially different conceptually in that norms are essentially normative, behaviour-independent and desire-independent, whereas conventions are not normative and are behaviour-dependent and desire-dependent. They are crucially different functionally in that, whereas conventions serve a coordination-facilitating function, we shall suggest that norms (including social norms) principally serve an *accountability*-creating function.

Getting clear on the conceptual and functional differences between norms and conventions is important, in the first instance, simply because it helps us to understand better their nature. A good way of understanding what something is is to

generally complied with) transform situations of conflict including cooperation problems *into* coordination problems.

⁴ The problem with the more ambitious program is that many if not most moral norms seem to arise in contexts where our interests are pitted against one another. In other words, they are solutions to cooperation rather than coordination problems.

⁵ Ullmann-Margalit (1977) holds an even more restricted thesis, namely, that *some* social norms are to be understood in terms of conventions.

focus on trying to say what it's *not*. But it is also potentially important for certain kinds of explanatory purposes. If we are right that norms are both more and less than conventions, the explanations of the emergence of norms that have been offered by proponents of the norms-as-conventions thesis are likely to turn out to be both inadequate in some ways and unnecessarily involved in others. Moreover, some theorists are sceptical that norms themselves are capable of playing any significant explanatory role over and above individuals' desires and non-normative beliefs. Once we realise the ways in which norms are more and less than conventions, it is possible that their explanatory significance goes rather further.

We shall begin by briefly rehearsing the notions of 'convention' and 'norm' that are at issue (section I). Next, we shall examine the conventions-as-norms thesis (section II) and the norms-as-conventions thesis (section III) in turn, arguing that both theses are mistaken. We shall then attempt to offer a functional characterisation of norms that can do something to make sense of the conceptual differences between norms and conventions and hence why we should expect them to come apart (section IV). In conclusion, we shall briefly mention several ways in which conventions may nonetheless be relevant to the *emergence* of norms (section V).

I. Preliminaries

Before we are in a position to investigate the relation between norms and conventions, we need to say something briefly about the notions of 'norm' and 'convention' that are at issue.

I.A. Conventions

Conventions in the sense that we are interested here are solutions to coordination problems. As we have already noted, we have no wish to deny that there are other concepts, some of which may be appropriately picked out by the word 'convention'. Thus, Margaret Gilbert (1989, 2008) uses the term 'convention' to mean something much more like 'norm'.⁶ Clearly, however, this won't help those who want to understand conventions *as* norms or norms *as* conventions. What makes the conventions-as-norms thesis and the norms-as-conventions thesis interesting for our purposes is that they purport to offer *accounts* of the one in terms of the other.

Conventions, then, are solutions to coordination problems – situations in which individuals' interests are aligned such that each desires that everyone behaves in the mutually advantageous way; and coordination is a problem because there are different ways in which they might do that. But what exactly does this mean? Lewis interprets it as follows:

A regularity R in the behavior of members of a population P when they are agents in a recurrent situation S is a convention if and only if it is true that, and it is common knowledge in P that, in almost any instance of S among members of P ,

- (1) almost everyone conforms to R ;
- (2) almost everyone expects almost everyone else to conform to R ;
- (3) almost everyone has approximately the same preferences regarding all possible combinations of actions;
- (4) almost everyone prefers that any one more conform to R , on condition that almost everyone conform to R ;
- (5) almost everyone would prefer that anyone conform to R' , on condition that almost everyone conform to R' ,

where R' is some possible regularity in the behavior of members of P in S , such that almost no one in almost any instance of S among members of P could conform both to R' and to R (Lewis 1969, 78).

It might seem, however, that some of Lewis's conditions are unnecessarily strong and that they should be relaxed or abandoned. Thus, some critics have argued

⁶ One important difference between our account and Gilbert's is that Gilbert explicitly abandons methodological individualism.

that Lewis' claim that conventions require universal or near universal conformity is too strong (Millikan 2005); that conventions need not be solutions to 'pure' coordination problems in the sense that there exist alternative, equally desirable solutions to the coordination problems to which they constitute solutions;⁷ that common knowledge at least in Lewis's sense of an infinite hierarchy of beliefs is not required (Gilbert 1989, ch. 3).

We shall assume that at least the following conditions are necessary. First, conventions are behavioural regularities and thus are *behaviour-dependent* in the sense that they require that individuals generally (though not universally or even near universally) behave in accordance with them. Second, conventions are *believed* regularities. It is not simply that individuals *in fact* generally behave in certain ways. Rather, individuals must generally *expect* that others will behave in those ways; they must have sufficient *confidence* that others will generally conform with the regularities that constitute the conventions. Third, conventions are *desire-dependent* regularities. People generally desire to conform with them conditional on others also generally conforming. Fourth, while conventions involve multiple equilibria, these do not have to be desired *equally* by different individuals. Finally, the presence of the relevant attitudes must at least partly *explain* the behavioural regularities. We shall remain neutral about whether conventions are, in addition, arbitrary and self-

⁷ This is captured by the idea that there exists an alternative regularity R' such that people prefer that everyone complies with R' (rather than not comply). Given Lewis's claim that individuals must have approximately the same preferences, this entails that individuals must prefer the same conventions. But some of Lewis's critics have suggested that this assumption should be abandoned. See e.g. Miller 2001. It is enough that individuals prefer that others comply with R than that they not comply (and perhaps that they prefer that individuals comply with R' than that they not comply). But this is consistent with some individuals preferring that everyone complies with R than with R' and others preferring that everyone complies with R'. That is to say, conventions are found even in Unequal Coordination Games like 'Battle of the Sexes'. Even this weaker assumption of arbitrariness has been challenged (e.g. by Sugden 1986, Marmor 1996, Vanderschraaf 1995).

reinforcing regularities, and about whether the fact that the relevant conditions obtain must be a matter of common knowledge.

I.B. Norms

So much for conventions. What about norms? There are two irrelevant senses of ‘norm’ that are worth mentioning in order to set aside. One is the purely *statistical* sense of ‘norm’ as simply denoting what is common or habitual, as when we say that ‘in Australian households, having two vehicles is the norm’. This need have no normative content whatsoever. We need not be implying that there is anything deficient about a household with more (or less) than two vehicles, that such a household ‘goes wrong’ in any way or is guilty of any kind of failing. We are simply describing what is common or habitual among Australian households. We are using the word ‘norm’ to make a statistical generalisation, asserting a claim with an implicit habitual quantifier.

A second sense of ‘norm’ is that in which it refers to an *objectively valid rule or normative principle*, irrespective of whether the principle is anywhere accepted or endorsed. Moral philosophers, for example, sometimes describe what they are doing as trying to identify ‘norms’ of morality or ‘norms’ of rationality. By this they mean that they are trying to identify objectively valid rules or normative principles of the relevant kind. Such principles may or may not be accepted by particular individuals or communities.⁸

The sense of ‘norm’ that is of concern to us in here is importantly different from both of these senses. We are interested in ‘norms’ understood, not as statistical

⁸ Consider, for example, the principle that ‘one must sacrifice one’s life if one can bring about more good by doing so’. When moral philosophers consider whether this is a ‘norm’ of morality, they are not asking whether the principle is accepted; their concern instead is simply to determine whether it is objectively valid.

generalisations or as objectively valid normative principles, but as something like *accepted rules or normative principles*. This is the sense that is at issue when we speak of the ‘norm’ that exists among the French of calling strangers ‘vous’ rather than ‘tu’; or the ‘norm’ that exists among the English that one must not stand too close to others when talking to them; or the norm that exists among members of the global community that states must respect one another’s territorial sovereignty. When we describe these as ‘norms’, what we are saying is that a particular normative principle is accepted in a particular group or community. Thus, for example, we are saying that there is an accepted normative principle among the French that one must call strangers ‘vous’ rather than ‘tu’, an accepted normative principle among the English that one mustn’t stand too close to others when talking to them, and an accepted normative principle among members of the global community that states must respect one another’s territorial sovereignty.

Norms in our sense therefore have two elements. First, they have a *normative* element. They involve and are partly constituted by *normative principles*. Wherever there is a norm, there will be a particular normative principle that the norm instantiates.⁹ Second, norms also have a crucial *socio-empirical* element. Like conventions, they are social facts. A normative principle is always a norm *in* or *of* a particular group or community. It makes no sense to inquire of any given principle – the principle that women must wear headscarves, say – whether it is a norm

⁹ Normative principles are *general requirements*. They are requirements inasmuch as they *require* things of agents; they describe what agents *must* and *mustn’t* do: refrain from murder; wear a headscarf; keep one’s promises; wear black at funerals; and so on. We can think of requirements as expressible either by deontic sentences ('S must X in C' and 'S mustn’t X in C') or by imperatives ('X in C!' and 'Don’t X in C!'). For convenience’s sake, we ourselves shall usually speak of them in terms of deontic sentences, but nothing of consequence turns on this. They are *general* requirements inasmuch as they are requirements that possess a certain generality of scope and application. They apply, not to the performance by a *particular* agent of a particular action in a particular circumstance, but to the performance by *any* agent (within a specified class) of certain types of acts in certain types of circumstances.

simpliciter. Norms are instead essentially tied to the particular groups or communities in which, or of which, they are norms. What makes a normative principle a norm of a particular group or community is that it is *accepted* in that group or community. Take the normative principle that one must pass the port to the left. What makes this a norm among Oxford dons is that among Oxford dons it is accepted that one must pass the port to the left.

We might wonder what it means to say that a normative principle is ‘accepted in a particular group or community’? Clearly, it isn’t necessary that the principle is *universally* accepted by the individual members of the group or community.¹⁰ What seems necessary, rather, is that a *significant proportion* of the members of the group accept the principle. This is a vague matter of course. But that’s a good thing. For it’s also a vague matter whether or not a given normative principle is a norm of a group.

What does it mean for an individual member of a group to ‘accept’ a normative principle? Clearly it is not enough that she simply *behaves* in the ways that the principle requires.¹¹ Virtually all Australians breathe, drink water, defecate. That does not mean that Australians accept normative principles that one breathe, drink water and defecate, respectively; that there exist norms in Australia to that effect.¹² Mere patterns of behaviour are not norms. That would be to divest norms of their normativity, to confuse them with ‘norms’ in the sense of mere statistical generalizations.

¹⁰ This would imply that if there were even a single dissident member of the group who did not accept the principle, then this would disqualify it from being a norm of that group. So if there were a single Oxford don who did not accept that one must pass the port to the left, then this would disqualify it from being a norm of Oxford dons. But that is patently absurd.

¹¹ The idea that norms are mere behavioural regularities has been defended by some theorists. For example, Peyton Young (2003, 390) writes that: ‘a norm is an equilibrium behavior in a game played repeatedly by many different individuals in a society where the behavior is known to be customary.’ Compare also Posner (2000, ch. 1).

¹² As Hart (1961, 9) notes: ‘Mere convergence in behaviour between members of a social group may exist (all may regularly drink tea at breakfast or go weekly to the cinema) and yet there may be no rule *requiring* it’.

Nor is accepting a normative principle a matter of *desiring* to act in the ways that the principle requires, either unconditionally or conditional on others also acting thus.¹³ So far as unconditional desires are concerned, presumably Australians desire unconditionally to breathe, drink water and defecate. Yet it would be very strange to say that for this reason they accept normative principles that require them to breathe, drink water and defecate. So far as conditional desires are concerned, we can imagine a closet lesbian strongly desiring to come out of the closet conditional on other lesbians also coming out of the closet without in any way thinking that she (or lesbians in general) *must* do so. It may be simply that the prospect of a world in which she didn't feel compelled to hide her sexual life strikes her as much more comfortable.

In order to be said to accept a principle, clearly something more is needed? What more? In his famous discussion of the differences between social rules and mere habits, H.L.A. Hart, suggested that the key is that accepting a rule or principle involves being prepared to take a 'reflective critical attitude' to certain behaviour in light of the principle. He illustrates this with the example of someone who accepts the rules of chess.

Chess players do not merely have similar habits of moving the Queen in the same way which an external observer, who knew nothing about their attitude to the moves which they make, could record. In addition, they have a reflective critical attitude to this pattern of behaviour: they regard it as a standard for all who play the game. Each ... 'has views' about the propriety of all moving the Queen in that way. These views are manifested in the criticism of others and demands for conformity made upon others when deviation is actual or threatened, and in the acknowledgement of the legitimacy of such criticism and demands when received from others. For the expression of such criticisms, demands, and acknowledgements a wide range of 'normative' language is used. 'I (You) ought not to have moved the Queen like that', 'I (You) must do that', 'That is right', 'That is wrong' (Hart 1961, 55-6).

¹³ This is how Bicchieri (2006) understands the notion of accepting a normative principle.

In short, accepting a normative principle involves certain kinds of *normative attitudes*: being disposed to regard certain responses as appropriate and inappropriate, the ‘right ones’ and the ‘wrong ones’; being disposed to criticize those who do what is wrong, and to regard such criticism as legitimate; and so on.

So much, then, for the notions of ‘norm’ and ‘convention’ that are at issue. Let us now consider attempts to identify a conceptual or constitutive relation between them. We shall begin in the next section by examining the conventions-as-norms thesis. We shall turn to the norms-as-conventions thesis in section III.

II. The conventions-as-norms thesis

The conventions-as-norms thesis holds that conventions are or at least entail special kinds of norms. Here is our argument against the conventions-as-norms thesis. First, norms, as we have seen, possess a certain kind of *normativity*; in particular, they entail particular normative principles that are accepted by a significant proportion of the members of the groups or communities in which they are norms. Social norms are no exception. Take the normative principle that women must wear headscarves in public. What makes this a social norm in one society (Saudi Arabia, say) and not a norm in another society (Scotland, say) is that in Saudi Arabia it is accepted by a significant proportion of Saudi Arabians that women must wear headscarves in public, and the fact of acceptance is a matter of common knowledge.¹⁴ But, second, conventions do not entail normative principles that are accepted by a significant

¹⁴ To appreciate the importance of common knowledge, suppose that it were to turn out that in our society an overwhelming majority of individuals accept that one must never, under any circumstances, engage in masturbation. But no one knows that others feel this way about masturbation; within popular culture, masturbation is presented as being not such a big deal. It would be odd, we contend, to say that there is a ‘norm’ within our society that one mustn’t engage in masturbation. In order for there to be a norm, it seems that the fact of acceptance must itself be a matter of common knowledge.

proportion of the members of the groups or communities in which they are conventions. To be sure, conventions are often *accompanied* by such normative principles. But the existence of a convention does not entail the existence of a generally accepted normative principle.¹⁵ Therefore, conventions are not, nor do they entail, corresponding norms.

Consider the following example. Suppose that a group of friends meets for lunch each Friday at noon at Imelda's Inn. They do so, let us suppose, because they are fond of one another and are eager to meet regularly in spite of living busy lives that do not afford them many opportunities for doing so. The current arrangement suits them all very nicely. Of course, other arrangements would have done just as well too; Imelda's is hardly the sole suitable establishment within easy walking distance of their respective workplaces; and 12:30 would have done just as well as 12:00. But nonetheless they stick to the current arrangement, knowing that it is mutually satisfactory one for all concerned.

This satisfies each of the conditions required for the existence of a convention. There is a regularity in the behaviour of a group of persons. The regularity is based on certain interdependent attitudes that are themselves a matter of common knowledge among the friends. Roughly, each wants to meet the others on a weekly basis at a suitable place at a suitable time, that is, to turn up at a suitable place at a suitable time conditional on the others also turning up at the same place at the same time. Each believes that the others will be at Imelda's, a suitable place, on Fridays at noon, a suitable time. So each desires that everyone (including himself) turn up at Imelda's on

¹⁵ This echoes Margaret Gilbert's argument against Lewis' conception of conventions (see Gilbert 1989, 2008). However, Gilbert assumes that social norms (which unfortunately she calls 'social conventions') are normative in a much stronger sense, namely, that a social norm is such that 'all equal, I ought to conform to it' (2008, 9). By contrast, we are concerned to show merely that they entail that individuals *accept* the requirements that constitute social norms.

Fridays at noon. (Each would have desired and taken himself to have a reason to turn up at some different place and time if the others had also turned up there and then.) Moreover, all of this is common knowledge among them. Each knows what the others believe and desire and each knows that the others know that they know it. In consequence, the regularity is self-sustaining; the friends generally turn up at Imelda's on Fridays at noon.

Must this convention involve a generally accepted normative principle requiring them to turn up, however? Lewis (1969, 97-100) says yes.¹⁶ Suppose, however, that each of the lunching friends has a peculiar loathing for normatively binding social arrangements – perhaps as a result of having been raised by moralistic and overly demanding parents. Thus, though they want to meet one another on a weekly basis, they only want to meet in this way, conditional on its being clear and explicit that there is *no requirement* to turn up and no criticism or censure that flows from their not doing so. Were they to feel subject to a requirement to turn up, or that the others took them to be subject to a requirement to turn up, or that the others took themselves to be subject to a requirement to turn up, any pleasure they might otherwise hope to enjoy from meeting would instantly disappear. Their meeting must

¹⁶ The crux of his argument is as follows. If there is a convention among our group G that members of G do X in circumstances C, this entails that, for most members of G, doing X in C will be (a) what we have reason to believe will answer to our own preferences and (b) what we have reason to believe will answer to the preferences of most of the other members of G and (c) what we have reason to believe most of the other members of G have reason to expect us to do (Lewis 1969, 97-8). According to Lewis, these are 'presumptive reasons why I ought to conform' with the convention (Lewis 1969, 98). This is enough to show that '[a]ny convention is, by definition, a norm [in the broad sense] which there is some presumption one ought to conform to' (Lewis 1969, 99). But Lewis thinks that he can also show that 'it is also, by definition, a socially enforced norm: one is expected to conform, and failing to conform tends to evoke unfavourable responses from others' (Lewis 1969, 99). This is because a convention in G that members of G do X in C also entails that one's doing X in C will be what most others members of G have reason to believe will answer to (a') one's own preferences, (b') what one has reason to believe their preferences are and (c') what one has reason to believe they have reason to expect one to do. 'So,' writes Lewis, if others see that we have failed to conform 'they will be surprised, and they will tend to explain my conduct discredibly. The poor opinions they form of me, and their reproaches, punishment, and distrust are the unfavourable responses I have evoked by my failure to conform to the convention' (Lewis 1969, 99).

be, so to speak, wholly optional on all sides. Moreover, all this is common knowledge among the friends. Given this common understanding, each of them duly shows up, believes that the others will turn up, and so on.

This is a case of a convention that does not involve the existence of a generally accepted normative principle requiring them to turn up. Indeed, it is a convention that *necessarily* involves the *non-existence* of such a principle, as a condition of the possibility of its own existence. Were there to be a generally accepted normative principle among the friends requiring them to turn up, the convention would cease to exist. It only exists insofar as it is not accompanied by such a principle. The problem with Lewis's argument is that it fails to take account of cases like this one where the 'presumptive reasons' he mentions are *necessarily non-obligating* from the perspective of the members of the group. It is simply not the case that the members of the group must accept that they are required to turn up. Nor is it the case that, were one of the friends to fail to turn up, the others 'will tend to explain [his] conduct discredibly'.¹⁷

Here, then, is the first conceptual difference between norms and conventions. Norms are necessarily normative. Conventions are not necessarily normative; some indeed are necessarily non-normative. This suffices to refute the conventions-as-norms thesis. Might the proponent of the conventions-as-norms thesis not respond by restricting the thesis to *some* conventions? After all, she might insist that we have merely shown that not *all* conventions are necessarily normative. For all that, perhaps *some* conventions are necessarily normative; and the existence of *these* conventions

¹⁷ This should be unsurprising. There are numerous cases where we have justified beliefs that others will act in ways that we know that they know we desire without taking them to be required to act in those ways. For discussion of a range of cases of this sort, see Owens (2006, 53-61) and Southwood and Friedrich (2009, 266-9).

may be sufficient for the existence of a corresponding norm. Perhaps this would suffice to save a more restricted version of the conventions-as-norms thesis.

This is based on a simple mistake. The lesson to be drawn from the case of Imelda's Inn is that for *any* convention, it is conceptually possible, given the psychology of the parties to the convention, that they do not accept a normative principle requiring them to comply. In no case do the conditions that are individually necessary and jointly sufficient for a convention *guarantee* that the parties to the convention must accept such a requirement. This is true even of conventions where we do, as a matter of fact, accept a corresponding normative principle, such as a drive-on-the-left convention. Even though we *happen* to conceive of driving behaviour in a normative way, there is no necessity to this. Our driving convention could survive even if we were all normative egoists who thought that facts about other individuals provide us with no reasons to act. Indeed, it could survive even if we were all to become normative nihilists who didn't think there was *anything* that we must do. Where conventions come with normative trappings, these trappings are not entailed by the existence of a convention as such; they are strictly extraneous. In the case of norms, by contrast, the normativity is essential to the kinds of entities they are.

III. The norms-as-conventions thesis

Let us now turn to the norms-as-conventions thesis. This holds that at least some norms are or entail special kinds of conventions. The norms-as-conventions thesis is explicitly restricted in two ways. First, it holds that the relevant class of norms are or entail *special* kinds of conventions. It is therefore not susceptible to the objection we presented against the conventions-as-norms thesis. A proponent of the norms-as-conventions thesis can concede the possibility of (essentially) non-normative

conventions, so long as she believes that such conventions fail to be special in the relevant sense. Second, the proponent of the norms-as-conventions need not claim that *all* norms are or entail special kinds of conventions, only that *some* do. We shall take the norms-as-conventions thesis to be a claim only about social or customary norms. As a claim about social or customary norms, it has considerable *prima facie* plausibility.¹⁸ Despite this, we shall argue that the norms-as-conventions thesis, too, is ultimately mistaken. We shall raise two objections to the norms-as-conventions thesis. The first concerns the behaviour-dependence of conventions. The second concerns their desire-dependence.

III.A. Behaviour-independence

As we saw above, conventions are *behaviour-dependent*. They are behavioural regularities; in order for them to be conventions, individuals must generally act in accordance with them.¹⁹ If motorists do not generally drive on the left in Australia –

¹⁸ First, it offers a straightforward explanation of what we might call the *customary* aspect of social norms. It seems to be a distinctive and important feature of social norms that they are somehow constitutively tied to the contingent social practices of the groups or communities in which they are norms. Conventions are, of course, paradigmatic instances of such social practices. The norms-as-conventions thesis can explain why this is so. Second, a philosophical characterisation of norms should ideally say something about their core *function*. Norms, like any social phenomenon, can be thought of as a kind of tool. An adequate understanding of any tool must presumably say something about what it's *for*. An account of what knives are that didn't say anything about cutting would be patently inadequate. The norms-as-conventions thesis offers an attractive account of the distinctive function of social norms. Conventions serve the function of facilitating coordination. This is clearly something that in many cases we have reason to want. Third, the norms-as-conventions thesis appears to give us a clear and operationalisable understanding of social norms that can form the basis for systematic social scientific investigation. Though the notion of a social norm may strike us as rather mysterious, the notion of a convention is relatively well understood, ontologically secure and explanatorily salient.

¹⁹ This is not to say that individuals *universally* (or even *almost universally*) act in accordance with them. As we saw above, Lewis (1969, 78) endorses the stronger claim. But, as a number of critics have persuasively argued, the stronger claim cannot reasonably be maintained. It is a convention in Australia that motorists drive on the left side of the road. Still, plenty of Australians can attest to encountering motorists driving on the right – either because they are drunk, or distracted, or American. Some Australians may even admit to having done so themselves on occasion. It is a convention in France that friends kiss when they greet one another. But the fact that they do not always do so does not undermine its status as a convention.

A referee objected that the claim that conventions entail that individuals generally act in accordance with them might still be too strong and should perhaps be abandoned in favour of the claim

if, for example, it is more or less random which side of the road they drive on – it is not a convention. If French friends do not generally kiss one another on the cheek when they greet, then, again, it can hardly amount to a convention. Norms, by contrast, including social norms, are *not* behaviour-dependent in this way. Of course it is perfectly true that *many* norms are, as a matter of fact, such that individuals generally act in accordance with them.²⁰ Nonetheless, it is not the case that norms are *necessarily* such that were individuals not generally to act in accordance with them, they would go out of existence. Therefore, social norms are not, nor do they entail, conventions.

Suppose, for example, that there is a social norm in Moldova to the effect that one mustn't urinate in public swimming pools (see Southwood 2010b). In order for this to be a social norm in Moldova, the principle that one mustn't urinate in public swimming pools must be generally accepted within the community. A sufficient number of Moldovans must accept the principle, where this involves having certain normative attitudes concerning urinating in public swimming pools: e.g. being disposed to think normative thoughts such as 'It's wrong to urinate in public swimming pools or and to disapprove of anyone who does so (including perhaps themselves). Moreover, this must be a matter of common knowledge. In other words, the Moldovans must generally know that other members of the Moldovan community generally accept that one mustn't urinate in public swimming pools, know that others know that they accept it, and so on. Does it follow that Moldovans must generally act

that conventions entail that individuals generally do not violate them. The referee suggested an example where cars cannot be used any more because of a massive oil shortage. Yet, according to the referee, the drive-on-the-left convention might still hold. We are skeptical about this example on the grounds that the *explanation* for why members of the society in question do not drive on the right no longer seems to have anything to do with their attitudes. Moreover, even in the society in question, there remains a regularity in behaviour: driving on the left insofar as one drives at all.

²⁰ We suspect this fact misleads many into supposing that that is a necessary feature of social norms. See Axelrod (1984) and Pettit (1990).

in accordance with the norm? Not at all. It is perfectly possible that urinating in public swimming pools is widespread among the Moldovans. To be sure, many of them will presumably experience a certain guilt or embarrassment whenever they do so; and if they ever happen to discover that others are doing it, to be suitably shocked and disapproving. They may, moreover, be disposed to approve of those who they believe refrain from urinating in public swimming pools. They may be disposed to encourage others from refraining from urinating in public swimming pools. They may take all these dispositions to be perfectly appropriate and warranted. For all that, urinating in public swimming pools may be rife.

What might the proponent of the norms-as-conventions thesis say in response? Perhaps it will be said that, while there is indeed a norm at play in the aforementioned example, we have misidentified its content. The true content of the normative principle that the Moldovans accept is not that one must refrain from urinating in public swimming pools but that one must disapprove of anyone who urinates in a public swimming pool. The requirement to disapprove of anyone who urinates in a public swimming pool *is* commonly complied with. So the example does not succeed in refuting the claim that social norms are behaviour-dependent.²¹

The problem is that this does not seem to come remotely close to capturing the content of the norm. What Moldovans principally object to is not others failing to disapprove of those who urinate in a public swimming pool. What they disapprove of is others urinating in a public swimming pool. Suppose that one member of the Moldovan community is caught urinating in a public swimming pool and another member of the community refuses publicly to disapprove of her. The other members of the community will disapprove fundamentally of the person who urinates in a

²¹ Thanks to Richard Bradley for discussion here.

public swimming pool, and any disapproval they display toward the non-disapprover will be parasitic upon that fundamental disapproval.

Or perhaps it will be said that, even if it is a norm of some kind, the principle that one mustn't urinate in public swimming pools cannot be a *social* norm because it fails to preserve the link that exists between the idea of a social norm and what is 'normal'. Not urinating in public swimming pools can hardly be the norm within a society unless it is 'normally the case' that people don't urinate in public swimming pools. But, by hypothesis, in the society in question – Moldova – urinating in public swimming pools is extremely widespread. Therefore, not urinating in public swimming pools cannot be the norm in Moldova.

This response trades on the ambiguity in the word 'norm' that we mentioned above. It is quite correct that the word 'norm' can be used to refer to statistical generalisations – to what is usual or habitual (as when we say that in western societies, 'teenage drinking is the norm', or that among politicians 'dishonesty is the norm', or that among elephants 'a single offspring is the norm'). But the sense of 'norm' that we are interested in is the sense in which it refers to a certain kind of accepted rule or normative principle. These are not two elements of the same concept. Instead, they are two fundamentally different concepts. To see this, it is helpful to note that, although we use the word 'norm' in both cases, we use quite different locutions to pick out the relevant concepts. In the case of 'norm' understood as accepted normative principle, we say that 'it is *a norm* that one X'. Whereas, in the case of 'norm' understood as statistical generalisation, we say that 'it is *the norm* that people X'. Moreover, the 'it is a norm' locution takes the subjunctive mood, whereas the 'it is the norm' locution takes the indicative mood. This can be easily missed in English when we use the third-person plural, as when say 'it is a/the norm that people

X'. But consider the third-person singular. We would say 'it is a norm in the Netherlands that *one bring* flowers or a bottle of wine for one's host when invited for dinner', whereas one would always say 'it is the norm in the Netherlands that *guests bring* flowers or a bottle of wine for one's host when invited for dinner' or, better, 'bringing flowers or a bottle of wine for one's host when one is invited for dinner is the norm in the Netherlands'.

Still, it does seem that social norms, unlike other kinds of socially accepted rules, have an important customary aspect. There is some kind of crucial conceptual link between the notion of a social norm and what is customary or habitual (Southwood 2010b; forthcoming). In particular, it seems that the generally held normative attitudes that are constitutive of social norms must somehow make reference to the way that people generally behave. Oxford dons judge that one must pass the port to the left in part *because* passing the port to the left is 'what is generally done' by Oxford dons. This is an important difference between social norms and *moral* norms. Consider the moral norm in Canada that one mustn't murder one's parents. Canadians do not judge that one mustn't murder one's parents *because* refraining from murder is what is generally done. Their judgements are independent of the way people generally behave. In the case of social norms, by contrast, it seems that the judgements must make essential reference to what is generally done. And the proponent of the norms-as-conventions thesis might insist that this is why the Moldovan norm is not a social norm. Hence, it doesn't constitute a counterexample to the norms-as-conventions thesis, suitably interpreted.

The question is what precisely the customary aspect of social norms amounts to. We have argued elsewhere that the normative attitudes that are constitutive of social norms must be *grounded*, at least in part, in presumed *social practices*, which

are a special kind of regularity of behaviour (Southwood 2010b, forthcoming).²² What this means is that in the minds of those who accept the relevant requirement the justification must come, at least in part, from a particular practice. So, for example, in the mind of the Oxford don, the practice of generally passing the port to the left must be part of what justifies the requirement to pass the port to the left. The point is that this is perfectly consistent with the don being *mistaken* about dons generally passing the port to the left. Perhaps the practice has disappeared or changed. Of course, since the act of passing the port happens to be a public one, it is not particularly plausible that dons could be *generally* mistaken about the port-passing practice. But there is nothing in the concept of a social norm as such to rule out the possibility of such widespread mistakes.

What does this mean for the Moldovans? In order for the norm forbidding urinating in public swimming pools to be a *social* norm, it must generally be the case that, in the mind of Moldovans, a general social practice of refraining from urinating in public swimming pools must be part of what justifies the requirement to refrain from urinating in public swimming pools. This could clearly be the case even if, in fact, there is no such practice, since, unbeknownst to most Moldovans, urinating in public swimming pools is widespread.

We have argued that whereas conventions are behaviour-dependent, norms are not necessarily behaviour-dependent. What this shows is that the existence of a corresponding convention is not *necessary* in order for there to be a norm or even a social norm. Norms (even social norms) do not entail conventions, even special normative ones.

²² More precisely, a social practice is a regularity in behaviour among the members of a group that is explained, in part, by the presence within the group of pro-attitudes (or beliefs about the presence of pro-attitudes) towards the relevant behaviour that are a matter of common knowledge.

Still, we might wonder whether there are *some* norms that are behaviour-dependent. Consider, for example, the norm that exists in many societies that one must pay one's taxes. It might be argued that this norm is behaviour-dependent, since people only take themselves to be required to pay taxes insofar as others are also generally paying their taxes.

To say that this makes the norm behaviour-dependent, however, is based on a misunderstanding. Rather, what is going on is that the *content* of the normative principle that individuals accept is one that makes a *conditional demand* – a demand conditional on the principle being such that individuals generally act in accordance with it. In other words, the principle that individuals accept is something like the following:

For any individual member *s* of society *S*, *s* must pay his or her taxes if but only if other members of *S* also generally pay their taxes.

Since the norm involves a principle that makes a conditional demand, it follows that insofar as others are not paying their taxes, individuals do not take themselves to be required to pay taxes. This does not mean, however, that the norm is itself behaviour-dependent. It would only be behaviour-dependent if, were individuals not generally to act in accordance with it, the norm would go out of existence. But, in fact, it is conceptually impossible for there to be general non-conformity with such a norm. The only way one can fail to act in accordance with it is to fail to pay one's taxes despite others generally paying their taxes. But if everyone is failing to pay their taxes, then it follows that one is no longer required oneself to do so. So, facts about how individuals are generally acting are particularly irrelevant in the case of such norms.

Are there other norms that are behaviour-dependent? What about norms that involve principles that make unconditional demands, such as the norm in France that one must call strangers ‘vous’ rather than ‘tu’? Suppose that French people were to begin to flout this norm *en masse*: to make a point of calling strangers ‘tu’ rather than ‘vous’. It is clear that the norm would not be long of this world.

It is certainly right that there is some kind of important *empirical connection* between how people behave and many norms. Given what we said above about the customary aspect of social norms, it also follows that there is an important conceptual connection between social norms and *beliefs* about how people behave. So, for example, if the ‘vous’ rather than ‘tu’ norm in France is a social norm, this entails that the French must commonly *believe* that French people commonly call strangers ‘vous’ rather than ‘tu’. But this does not mean that the French must *in fact* act in accordance with the norm. Of course, typically, where there is a widespread failure to do so, this will come to be a matter of common knowledge. But the mere fact that the French stop behaving in a certain way is not enough, in and of itself, to entail the disappearance of the norm. Conventions, by contrast, are behavioural regularities; if there is no behavioural regularity, that is enough for there to be no convention, whatever people happen to believe about how others behave.

It seems, then, that we should conclude not merely that it is not the case that all norms are behaviour-dependent, but that all norms are behaviour-*independent* in the sense that they are consistent with a widespread failure to act in accordance with them. This is not a trivial feature of norms. On the contrary, it seems to be an important aspect of their normativity. For it seems that there is a certain kind of

normativity that entails the possibility of failure.²³ It has to be possible to fail to act in accordance with them (perhaps as well as possible to succeed in acting in accordance with them). The intuitive idea is that norms involve demands on one; and demands are the kinds of things that one can fail to live up to. Perhaps ‘ought’ implies ‘can’, but it doesn’t imply ‘does’. That would be to milk the norm of its normative oomph. The idea is usually understood as applying to particular individuals. But there is some temptation to think that it may also have application to groups of individuals. There seems to be something excessively normatively fragile about the idea of a demand that is guaranteed to go out of existence simply on account of individuals generally failing to act in accordance with it. That would not be a demand worthy of the name.

III.B. Desire-independence

Now we come to the other problem with the norms-as-conventions thesis. Conventions, as we have seen, are *desire-dependent* in the sense that parties to conventions must generally desire to act in accordance with the regularities that constitute conventions conditional on others also doing so. That is to say that *all-things-considered* they prefer to act in accordance with the regularity than not to do so, conditional on other also doing so.²⁴ Norms, by contrast, including social norms, are not desire-dependent. It is not the case that individuals must desire to act in accordance with norms conditional on others also doing so. So norms are not – nor do they entail – conventions.

²³ For critical discussion see Lavin (2004).

²⁴ The ‘all-things-considered’ is important in order to allow for conventions that are solutions to impure coordination problems.

To see that social norms do not entail desire-dependence, we might recall the case of the Moldovans.²⁵ However, we shall set aside social norms with which individuals do not generally act in accordance. Take instead the social norm that exists in Australia that people not go completely naked on the beach. In order for this to be a social norm, must it be the case that Australians generally prefer all-things-considered to refrain from total nudity on the beach than not to refrain, conditional on others also refraining from total nudity on the beach? Surely not. It may be that all Australians secretly have an overwhelming desire that everyone goes naked on the beach. Indeed, it may even be that the prevalence of the desire is somehow *due* to the existence of the norm. The proscription against nudity on the beach may be part of what explains its secret appeal. What is necessary for the nudity norm to exist is that Australians generally have whatever attitudes are constitutive of accepting the normative principle that holds that one must not go completely naked on the beach – at a minimum, that they are disposed to think normative thoughts like ‘I must not go naked’ and to be disapproving of anyone who does so. But clearly they can have these attitudes without desiring to refrain from nudity, conditional on others also refraining from nudity.²⁶

²⁵ It’s simply not the case that members of the Moldovan society must desire to refrain from urinating in public swimming pools conditional on others also refraining from urinating in public swimming pools. They may or may not desire that others refrain from urinating in public swimming pools. But they certainly don’t desire to refrain from urinating in public swimming pools themselves, whether or not others refrain from urinating in public swimming pools. They may have some higher-order desire concerning urinating in public swimming pools; that is, perhaps they desire not to desire to refrain from urinating in public swimming pools. But patently they need not desire to refrain from urinating in public swimming pools. What is necessary, rather, is that they accept the normative principle that one must not urinate in public swimming pools. That plausibly involves various attitudes such as the disposition to disapprove of anyone (perhaps including themselves) who urinates in public swimming pools. But it need not involve having the desire to refrain from urinating in public swimming pools conditional on others also refraining from urinating in public swimming pools.

²⁶ What about the claim that ‘having a desire to X’ just is ‘being motivated to X, given one’s beliefs’? This is not a plausible position. The kind of dispositionalism about desires that it involves, while popular in the heyday of logical behaviourism, has been conclusively refuted many times over. See e.g. Geach (1957), Putnam (1965) and Strawson (1994). This is not to deny that there are more complex forms of dispositionalism that are considerably more plausible. However, they are more

One possible response at this point would be to insist that, insofar as the individuals in the examples really do fail to desire to act in accordance with the normative principles conditional on others doing likewise, this means that they don't really accept those principles. They may *say* they accept the principles. They may falsely believe that *others* accept the principles. But they don't really accept the principles themselves.

But why should we think that accepting a normative principle requires desiring to act in accordance with that requirement conditional on others acting likewise? Surely the right thing to say is that some of the principles we accept are ones that we desire to act in accordance with and others are ones that we don't. It's convenient and pleasant when these line up, of course, but it's hardly essential. There are all kinds of normative principles that we accept, the fulfilment of which is onerous, awkward or even downright unpleasant. To insist that the only principles we can accept are those with which we desire to act in accordance, conditional on others also acting in accordance with them, seems to make life much easier and simpler than we know it to be.²⁷

Might the proponent of the norms-as-conventions thesis appeal to the internalist thesis that an internal connection exists between accepting a normative principle and being motivated to act? Thus, to accept a normative principle that holds

plausible precisely because they deny that desiring to X entails being actually motivated to act. Moreover, at best it succeeds in saving the norms-as-conventions thesis from the nudity case. It does not help in the Moldovan case. Indeed, if desiring just is being disposed to act, then it follows straightforwardly that urinators don't desire to refrain from urinating in public swimming pools.

²⁷ This is not to deny that normative attitudes and desires may be interestingly linked in some ways. Thus, for example, it may be that it is not psychologically possible to accept a normative principle that requires us to act in ways to which we have an *overwhelming aversion*. Moreover, it may be that we are *rationally required* to desire to act in accordance with *all-things-considered normative principle* that we accept. But to insist that it is literally impossible to accept a normative principle that requires one to X and yet fail to desire to X conditional on others also Xing seems simply too strong. It seems utterly false to the phenomenology of the normative thought that underpins and constitutes social norms.

that one must give money to charity might be thought to involve some disposition to give money to charity. If we have no such disposition to give money to charity, it is tempting to think that we have not made a genuine normative judgement, but only an 'as if' normative judgement, an 'inverted commas' normative judgement.

Many deny there is any internal connection between accepting a normative principle and motivation (Foot 1972, Brink 1989, Roskies 2003). But let us grant the existence of such a connection for the sake of argument. Does this mean that Australians don't really accept the requirement to refrain from total nudity on the beach? It would only do so if we were to understand the internal connection between normative judgement and motivation in a certain way, such that accepting a normative principle entails having an *all-things-considered preference* to act in accordance with the principle than not to do so. But this is not a kind of internalism that many will find attractive. It would mean, in effect, that we are always guaranteed to prefer to comply with than to violate any normative principle that we accept. More commonly, internalists understand the connection between normative judgement and motivation in quite a different way. Some hold that accepting a principle entails having some (albeit *defeasible*) desire to act in accordance with the requirement (Falk 1952). Others hold that accepting a principle entails being motivated to act in accordance with the requirement insofar as one is *rational* (Nagel 1970, Korsgaard 1986, Smith 1994). These weaker internalist theses are considerably more plausible. But, of course, the case of the Australian beachgoers is perfectly consistent with them. Thus, it may be that, although the Australians would prefer to go totally naked on the beach, they have some desire not to. Or again, it may be that, in preferring total nudity, they are to some extent irrational. In sum, the only kind of internalism that the case of the

Australian beachgoers is not consistent with is one we have good independent reasons to reject.

A different response would be to insist that, even if accepting a normative principle that requires one to X doesn't entail desiring to X conditional on others also Xing, the existence of a *social norm* that requires one to X *does* entail that individuals generally desire to X conditional on others also Xing, because of a generally held desire to act in accordance with others expectations of us or to avoid disapproval.²⁸ Why think that? Well, accepting a normative principle plausibly does involve having expectations of others to the effect that they will act in accordance with the requirement, and being disposed to disapprove of anyone who violates the requirement. Moreover, it is plausible to suppose that the fact that individuals generally accept the principle must be common knowledge. So let us suppose that social norms entail that individuals know (or believe) that others will expect them to act in accordance with the norms and disapprove of them for failing to do so. If we make the assumption that individuals generally desire to act in accordance with others expectations of them or to avoid disapproval, it seems to follow that where there is a social norm, individuals will generally desire to act in accordance with the norm conditional on enough others also doing so.²⁹

But even if we concede for the sake of argument that individuals generally have *some desire* to act in accordance with social norms conditional on others also so

²⁸ This appears to be Bicchieri's argument in Bicchieri (2006, ch. 1).

²⁹ One thing to say at this point is that even if this is correct, it does not show that social norms *entail* desire-dependence. Presumably even if, in the actual world, individuals generally desire to act in accordance with others' expectations of them or to avoid disapproval, we can easily imagine worlds where this desire is absent. Still, it might be said that the existence of the desire in the actual world casts serious doubt on our counterexamples. Since it is reasonable to assume that the individuals in these cases have the desire, we were wrong to suppose that they lacked the desire to act in accordance with the norms. Moreover, it might be added that a world in which we lacked the desire would be sufficiently distant to the actual world that it becomes doubtful whether social norms would have an important place there.

acting, it does not follow that the individuals prefer to act in accordance with the norms rather than not to conditional on others complying. Take the Moldovans. Let's concede that the fact (if it is a fact) that the Moldovans have a desire to act in accordance with others expectations of them or to avoid disapproval means that the urinators have some desire not to urinate in public swimming pools. But this doesn't mean that they prefer not to urinate in public swimming pools rather than to urinate conditional on others not urinating. On the contrary, whether or not others urinate, they may perfectly well prefer to urinate than not to urinate, taking all their desires, including the desire to avoid disapproval, into account.

Here, then, we have the third conceptual difference between norms and conventions. Whereas conventions are essentially desire-dependent, it is not the case that norms are essentially desire-dependent. Once again, however, we may wonder whether *some* norms (perhaps some *social* norms, say) are nonetheless desire-dependent in the sense that their continued existence is conceptually incompatible with individuals not generally desiring to act in accordance with them, conditional on others also acting in accordance with them.

We are inclined to think not. For any norm, it seems conceptually possible that the individuals who accept the relevant normative principle do so without desiring to act in accordance with the principle conditional on others also doing so; their acceptance of the principle is not determined simply by whether or not they have the relevant desire. Take the norm requiring individuals to wear black at funerals. We can surely imagine individuals coming to detest the idea of wearing black at funerals and yet continuing to accept that they must do so.

In saying this, we are not denying that there is an important *correlation* between the normative principles we accept and our desires to act in accordance with

them. Clearly we do, as a matter of fact, desire to conform with many of the norms we accept, at least insofar as we take others to comply in turn. The question is what is the status of this correlation. We suspect that the character of our desires is often the result of the normative attitudes we happen to accept, and that changes in the former are often due to changes in the latter. Our desires often come to reflect our normative convictions. But we have no wish to deny that the reverse may also be true; that facts about individuals' desires may have a causal impact on what normative principles are accepted and hence on what norms exist. Moreover, individuals may obviously take desires (their own or those of others) to be of *normative relevance*. Thus, changes in desires may occasion a shift in what they take ourselves to be required to do. The mechanism here is not a casual one so much as one that involves the application of normative principles to a new empirical situation. What seems hard to countenance is that there are norms for which conditional desires are a conceptual prerequisite.

In short, then, it seems not merely that norms differ from conventions in that it is not the case that all norms are desire-dependent. Rather, it seems that all norms are desire-independent in the sense that their continued existence is consistent, at least conceptually, with it not being true that individuals generally desire to act in accordance with them, conditional on others also doing so.

IV. The function of norms

What has emerged from our discussion is that norms (including social norms) and conventions are very different (even if often overlapping) kinds of social phenomena. First, norms are *normative* in the sense that they involve generally accepted requirements, whereas conventions are not necessarily normative. Second, norms are *behaviour-independent* in the sense that individuals need not generally act in

accordance with them, whereas conventions are behaviour-dependent. Third, norms are *desire-independent* in the sense that individuals need not generally desire to act in accordance with them conditional on others also doing so, whereas conventions are desire-dependent.

At this point, however, one may have the following reaction: Even if we are right, why is the category of norms that we have singled out an interesting and important one? Inasmuch as they serve a core coordinating function, conventions are directly and straightforwardly connected to things we have reason to want. In contrast, norms, as we have portrayed them, may seem to lack an obvious function.³⁰

Certainly, they do not appear principally to serve a coordinating function. To be sure, many norms surely do, as a matter of fact, help in facilitating coordination. Norms clearly serve many valuable purposes. However, we are looking for their *core function*. And it is not plausible to suppose that their core function is to facilitate coordination. To see this, suppose that we are operating in interactional contexts where coordination is either impossible or unimportant. It may be impossible because of some kind of irreconcilable conflict of interest, or because there is a sufficiently strong incentive to defect, or because the cost of acting in the interests of others is simply too great. It may be unimportant because we are already coordinating perfectly adequately as is – suppose, if you like, that there are plentiful stocks of affection among us – or because it is a situation in which there is simply no need to coordinate.

³⁰ Thinking about the functions of norms and conventions – what they *do* – affords us a deeper understanding of what they *are*. We can think of norms and conventions as certain kinds of tools. An adequate understanding of any tool must presumably say something about its core functional properties. An account of knives that didn't say anything about cutting would be patently inadequate. What seems true of knives seems equally true of norms and conventions. Thinking about the functions of norms and conventions *may* also turn out to be directly or indirectly relevant to the issue of explaining why they arise. To be sure, this is to enter dangerous territory, territory that we have been warned off by a generation of philosophers of social science (Elster 1982). We shall not pursue the issue here.

Under these circumstances, might there remain an important role for norms? The answer seems obviously to be 'yes'. Many of the most significant norms – including social norms – exist precisely in such situations. Think of norms requiring us to perform especially individually costly actions. Or think of norms compliance with which simply signals our sense of identification with the group. Or think of norms with which we express our shared values. If we are thinking of norms as a tool for serving our coordination interest, these kinds of norms must strike us as idiosyncratic outliers. But that seems quite mistaken. They seem, on the contrary, to be paradigmatic instances of the phenomenon.

Another way to bring this out is to note that the respects in which norms differ from conventions may seem to make them into a rather odd coordinating device, one that is unnecessarily complex in some ways and insufficiently connected to our interests in others. Take the *normativity* of norms. As we have seen, norms are normative in the sense that they entail generally accepted requirements. From the perspective of our coordination interest, this seems otiose. Of course, the normativity of norms may *sometimes* be helpful in ensuring stable coordination. However, conventions, as we saw, are not necessarily normative; and this fact does not necessarily undermine their ability to facilitate coordination. So normativity seems to be something of an explanatory dangler; it seems to be overreaching what is strictly required. Next, consider the *behaviour-independence* of norms. Norms are not necessarily undermined if those who are subject to them fail persistently to act in accordance with them. But, as we saw, effective coordination requires precisely that individuals generally behave in ways that others expect them to behave. Or again, consider the *desire-independence* of norms. This means that it is possible that individuals do not generally desire to act in accordance with norms conditional on

others also doing so. But coordination is only possible in cases where individuals' interests are aligned in ways that make certain forms of interaction mutually beneficial.

If norms don't principally serve a coordination function, then what kind of function *do* they serve? We suggest that their core function is to make us *accountable* to one another.³¹ That is, they create a generally recognised right or entitlement to demand and expect things of one another. Being accountable to one another in this sense is quite different from possessing the kind of reliable information about how other individuals will act that is an essential ingredient in effective coordination. What it involves is having a recognised right or entitlement to determine how others are to behave. When we become accountable to one another, we effect a normatively significant modification in our relations with each other. It is not that we have reliable information about what others will do. Rather, we are in a position to hold one another to account; to demand and expect things of one another.

Why might being accountable to one another be something that we have reason to want? Part of the answer is simply that there are many contexts in which others' actions adversely affect us and yet we cannot rely on them to take adequate account of our interests. It sometimes happens that our interests are fundamentally in competition and there is no special relationship between us that would act as a fetter to our unbridled pursuit of self-interest. It may be simply too tempting to commandeer another's possessions, to seduce a colleague's attractive partner, to dump one's litter on a neighbour's lawn, to leave one's coffee cups unwashed in the office kitchen, and so on. In short, there is a serious risk of our running roughshod over one another,

³¹ For a recent account of accountability, see Darwall (2006). See also Southwood (2010a, ch. 5).

indeed of our doing so in a wholly rational way. What each of us has reason to want, in such cases, is a way of limiting the right of each individual to make decisions in a purely self-regarding fashion (Checkel 1999). An obvious way to do this is for each individual to cede some degree of authority, so long as others also cede authority in turn.³²³³

But the accountability that norms create also plays other roles in social life. Norms can imbue behaviour with social meaning. They thereby enable us to express shared values, meanings and identities. The accountability-creating nature of norms plays a particular role in this expressive domain. An important theme in recent jurisprudence emphasises the expressive power of legal norms in particular to create meaningful categories, to valorise (and devalorise) certain forms of conduct and ways of life (Sunstein 1996, Cooter 1998, McAdams 2000, Anderson and Pildes 2000). This expressive power derives, in part, from the fact that law is authoritative; where a law forbids a certain kind of activity, this amounts to an authoritative pronouncement on the legitimacy of that activity. But the same is also true of non-legal norms. Where we have a shared right to demand things of one another, this affords us comparable opportunities to pronounce authoritatively on various dimensions of social life. In doing so, we project a certain image and craft a sense of self: of people who care about the dignity of human beings, for whom death is a sombre and serious business, for whom the group is more important than the individual, or whatever. We are in effect expressing to one another, through our recognition of one another's right to

³² Another possibility is of course to cede authority to the state.

³³ Indeed, arguably this is the best way, not merely of protecting our interests in general, but of safeguarding our interest in having the right to make individual decisions. As David Owens (2006, 70) nicely puts it, our interest in being individually authoritative may be 'best served by a bit of give-and-take'.

demand this or that, what matters to us, who we take ourselves to be, how we see ourselves and others.

It is not hard to see why norms should be thought to be perfect tools for creating accountability. Consider, first, the normativity of norms. For a norm to be in existence, it is not enough that most people think, 'I shall do this' or even 'I shall do this because I desire to do so (given my beliefs about others' beliefs, desires and behaviour)'. Rather, they must think in some sense, 'I am *required* to do this'. This is not to say that the normative thoughts must be in the forefront of their mind at every turn. Just as we manage to believe that the sun will rise tomorrow, that Canberra is in Australia and that Nicolas Sarkozy is the current French president without much in the way of occurrent thought, so too we manage to accept norms (that require us to put our rubbish in the bin, not to stand too closely to people, to eat lunch with our knives and forks in the correct hands) without thinking constantly, 'This is what I must do'. Nonetheless, norms entail that a sufficient number of people *accept* the normative principles that figure in them. When we accept a normative principle, we accept the right of others to expect us to comply with the principle. In doing so, we necessarily regard ourselves as accountable to others so far as complying with the principle is concerned.

Next, take the behaviour-independence and desire-independence of norms. One thing to say here is that behaviour-dependence and desire-dependence are not *needed* for accountability. Being able to hold others to account involves less as well as more than possessing the kind of reliable information about how they will act and what their interests are that is an essential ingredient in effective coordination. But we can go further. Not only are behaviour-dependence and desire-dependence not needed for accountability. They seem to be *inconsistent* with it. Consider a parent whose

commands are recognised by the child as legitimate but only because the parent always commands the child to do only what he knows the child will do or desires to do anyway. This would not be a parent with the ability to hold her child to account. To make accountability dependent on behaviour and desire in this way would be to milk it of all of its normative oomph in a way that would seem to be deeply antithetical to the very idea.

Norms, then, are perfectly suited to the business of creating accountability. Conventions obviously aren't. Accountability is an out-and-out normative notion. Any social institution, such as a convention, that does not necessarily possess this in-built normativity will be ill-suited to realising it. Similarly, the fact that conventions are behaviour-dependent and desire-dependent means that they do not necessarily possess the kind of normative oomph that we just saw was required for genuine accountability.

It is worth noting, finally, that a number of norms that seem rather baffling from the perspective of our coordination interest seem anything but baffling from the perspective of our accountability interest. Consider, first, norms that exist in spite of the fact that our interests are pitted against one another and hence coordination is impossible, norms where we would like others to act in accordance with them but would prefer not to do so ourselves. Or consider norms that arise in familial or other contexts where coordination is either not necessary (perhaps because we are already coordinating perfectly well as is or because coordination does not seem important in the contexts in question) or at any rate where coordination is not being achieved by the norms (perhaps because the norms are being persistently violated, such as the urinating norm we discussed above).

That such norms exist is perfectly comprehensible according to the accountability account that we have sketched. What all these norms do is create a certain kind of accountability to others. In virtue of the existence of a norm, we come to have a recognised right to demand that others refrain from littering, or pass the port to the left, or reciprocate acts of generosity, or whatever. Such accountability may be valuable in a number of ways, as we saw above. Norms thereby operate as vehicles for achieving things that we have reason to want: to compensate for deficiencies in altruism, or to create and galvanise meanings and identities, or to express shared values and to create new kinds of relationships and communities with our fellows.

Of course, to say that norms serve the function of creating accountability is not to say that we have reason to be particularly thrilled about each and every instance where such accountability is created. There is no shortage of horrendous norms: norms that are discriminatory, or unjust, or just plain stupid. Norms are a tool that serves a special function; and like any tool, they can be abused. Knives serve the function of cutting. Unfortunately, this means that knives can be used to kill as well as to cut bread and cheese. Conventions serve the function of facilitating coordination. Unfortunately, this means that they can be used to enact evil joint ventures as well as noble ones. Norms are just the same. Norms serve the function of creating accountability. This can be used to encourage ridiculous forms of self-sacrifice (think of the norms to which zealots appeal in order to mobilise suicide bombers), or to create and foster bigoted categories and meanings and express awful values (think of racist or misogynistic norms), or to constitute repressive, hierarchical, exclusive and unjust relationships (think of norms that help define and constitute highly unequal and patriarchal relationships, or those between masters and slaves and bullies and victims). Conceding these plain and undeniable truths should do nothing to undermine

our confidence in the claim that social norms serve the function of creating accountability, any more than conceding that cutting implements can be used to kill should undermine our confidence in the claim that knives serve the function of cutting. Creating accountability is simply what norms do.

V. The emergence of norms from conventions

We have argued that norms (including social norms) and conventions are crucially different conceptually and functionally, such that it is a serious mistake to assimilate them either. In saying this, it is important to emphasise that we are not, of course, denying that many norms may *emerge* from conventions. According to the account we have sketched, this would involve a process where intrinsically non-normative, desire-dependent behavioural regularities that serve a coordination-facilitating function metamorphose into something essentially normative, behaviour-independent and desire-independent that serve an accountability-creating function. The question of how that occurs is a fascinating one that has received extensive treatment elsewhere (see Lewis 1969, Ulmann-Margalit 1977, Sugden 1986, Coleman 1990, Bicchieri 2006). Here we shall just restrict ourselves to several brief observations.

First, notice that convention-driven activity may often fall under the jurisdiction of more general norms. There are many ways in which this might happen. The effects that a particular behaviour has on others can change significantly once the behaviour in question has become a convention. Driving on the right after the drive-on-the-left convention has been established is simply not the same action as driving on the right beforehand. Now it means deliberately risking an accident, which, in turn, brings into play certain general accountability-conferring norms such as ‘do not unnecessarily risk other people’s lives’ and ‘show consideration for others’. Moreover, by

participating repeatedly in a cooperative activity, we may tacitly signal to others that we *consent* to their having certain *expectations* of us. Finally, repeated convention-governed coordinated interaction can result in others coming to *rely* (and in our knowing that they rely) on the continuation of such a pattern in their planning and decision making, and thereby incurring potentially significant costs if the activity ceases.

Second, although being able to hold others to account is not strictly speaking *necessary* for coordination, it may often help. For one, even though a convention, by definition, should not involve any temptation to violate it, in real life violations do occur. People may make mistakes. They may act irrationally. Or the convention may interact adversely with other activities in which they are simultaneously engaged. Much coordination concerns behaviour in which there is something important at stake for us: our personal safety, for example, in the case of driving. If failure to comply would seriously impact upon others, they may acquire an interest in having some say over how others behave. If this interest can be channeled in such a way as to bring about what is needed, an accountability-creating norm may thus arise, as it were, in support of the coordination-facilitating convention.

Accountability may also help preserve the conditions that guarantee the possibility of sustained coordination over time. Individuals' desires, the weights of those desires and their willingness to act on them may all vary fairly dramatically, and be known by others to vary fairly dramatically. Yet coordination, if it is to be anything more than fleeting, requires considerable known stability in the attitudes that support the regularities that make it possible. Here is where accountability may be able to help. We are creatures for whom accountability appears to be a motivationally powerful force. The presence of an accountability-creating norm corresponding to a

convention may therefore help to achieve a kind of stability in individuals' desires that may be otherwise difficult to achieve. Desires grounded purely in self-interest may be fickle, varying in strength and motivational efficacy. It is not so (or anyway much less so) for desires grounded in a concern with the normative expectations we know others have of us. Those desires may have – and be known to have – a durability that helps facilitate successful coordination over time.

Finally, conventions seem to play a particularly important role in the emergence of *social* norms. The account of social norms we sketched above helps to explain why. What is distinctive of social norms is that the normative attitudes constitutive of social norms are grounded, at least in part, in presumed social practices. Conventions are of course paradigmatic social practices. Often, what seems to happen in the case of the emergence of social norms is that a particular kind of intrinsically meaningless behaviour – passing port to the left, wearing black at funerals, or whatever – acquires normative significance in people's mind through familiarity and simple habituation.

Obviously, providing a comprehensive account of the dynamics of norm emergence lies well beyond the ambitions of the current essay. Our aim here has been simply to get clear on the conceptual and functional differences between norms and conventions.³⁴

References

³⁴ We are grateful to Joel Anderson, Michael Baurmann, Geoff Brennan, Daniel Friedrich, Bob Goodin, Stephan Hartmann, Pauline Kleingeld, Eric Schliesser, Herman Siemans, Kai Spiekermann, Bruno Verbeek, audiences at the Universities of Leiden and Tilburg, and two anonymous referees for this journal for helpful feedback on earlier drafts. Research for the article was carried out under ARC Discovery Grant DP0663060.

- Anderson, E., and R. Pildes. 2000. Expressive theories of law: a general restatement. *University of Pennsylvania Law Review* 148: 1503-75.
- Axelrod, R.. 1984. *The evolution of cooperation*. New York: Basic Books.
- Bicchieri, C. 2006. *The grammar of society: the nature and dynamics of social norms*. New York: Cambridge University Press.
- Brink, D. 1989. *Moral realism and the foundations of ethics*. Cambridge; New York: Cambridge University Press.
- Checkel, J. 1999. Norms, institutions, and national identity in contemporary Europe. *International Studies Quarterly* 43: 83-114.
- Coleman, J. 1990. *Foundations of social theory*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Cooter, R. 1998. Expressive law and economics. *Journal of Legal Studies* 27: 585-608.
- Darwall, S. 2006 *The second-person standpoint: respect, morality, and accountability*. Cambridge, Mass.: Harvard University Press.
- Elster, J. 1982. Marxism, functionalism, and game theory: the case for methodological individualism. *Theory and Society* 11: 453-482.
- Falk, W.D. 1952. 'Ought' and 'motivation'. In *Readings in ethical theory*, ed. W. Sellars and J. Hospers. New York: Appleton-Century-Crofts.
- Foot, P. 1972. Morality as a system of hypothetical imperatives. *Philosophical Review* 81: 305-16.
- Geach, P. 1957. *Mental acts: their content and their objects*. London: Routledge and Kegan Paul.
- Gilbert, M. 1989. *On social facts*. London; New York: Routledge.
- Gilbert, M. 2008. Social convention revisited. *Topoi* 27: 5-16.
- Hart, H. L. A. 1961. *The concept of law*. Oxford: Clarendon Press. Korsgaard, C. 1986. Skepticism about practical reason. *Journal of Philosophy* 83: 5-25.
- Lavin, D. 2004. Practical reasons and the possibility of error. *Ethics* 114 (3): 424-57
- Lewis, D. 1969. *Convention: a philosophical study*. Cambridge, Mass.: Harvard University Press.
- McAdams, R. 2000. A focal point theory of expressive law. *Virginia Law Review* 86: 1650-1729.
- Marmor, A. 1996. On convention. *Synthese* 107: 349-371.
- Miller, S. 2001. *Social action: a teleological account*. Cambridge: Cambridge University Press.
- Millikan, R. 2005. *Language: a biological model*. Oxford: Clarendon Press.
- Nagel, T. 1970. *The possibility of altruism*. Oxford: Clarendon Press.
- Owens, D. 2006. A simple theory of promising. *Philosophical Review* 115: 51-77.
- Pettit, P. 1990. Virtus normativa. *Ethics* 100: 725-755.
- Posner, E. Symbols, signals, and social norms in politics and the law. *Journal of Legal Studies* 27: 765-98.
- Posner, E. 2000. *Law and social norms*. Cambridge: Harvard University Press.
- Putnam, R. 1965. Brains and behavior. In *Analytical philosophy*, ed. R.J. Butler. Oxford: Blackwell.
- Roskies, A. 2003. Are ethical judgments intrinsically motivational? Lessons from acquired sociopathy. *Philosophical Psychology* 16: 51-66
- Smith, M. 1994. *The moral problem*. Oxford; Cambridge, Mass.: Blackwell.
- Southwood, N. 2010a. *Contractualism and the foundations of morality*. Oxford: Oxford University Press.
- Southwood, N. 2010b. The authority of social norms. In *New waves in meta-ethics*, ed. M. Brady. Palgrave MacMillan: 234-48.

- Southwood, N. Forthcoming. The moral/conventional distinction. *Mind*.
- Southwood, N., and D. Friedrich. 2009. Promises beyond assurance. *Philosophical Studies* 144: 261-80.
- Strawson, G. 1994. *Mental reality*. Cambridge, Mass.: MIT Press.
- Sugden, R. 1986. *The economics of rights, co-operation, and welfare*. New York: Palgrave Macmillan.
- Sunstein, C. 1996. On the expressive function of the law. *University of Pennsylvania Law Review* 144: 2021-2031.
- Ullmann-Margalit, E. 1977. *The emergence of norms*. Oxford: Clarendon Press. 1977.
- Vanderschraaf, P. 1995. Convention as correlated equilibrium. *Erkenntnis* 42: 65–87.
- Verbeek, B. 2002. *Instrumental rationality and moral philosophy: an essay on the virtues of cooperation*. Dordrecht; Boston: Kluwer Academic Publishers.
- Voss, T. 2001. Game theoretical perspectives on the emergence of social norms. In *Social norms*, ed. M. Heidder and K.D. Opp. New York: Russel Sage: 105-36.
- Young, P. 1998. *Individual strategy and social structure*. Princeton, NJ: Princeton University Press.
- Young, P. 2003. The power of norms. In *Genetic and cultural evolution of cooperation*, ed. P. Hammerstein. Cambridge, Mass.: MIT Press: 389-99.