



2.2 | Can Evolutionary Psychology Assist Logicians? A Reply to Mallon

Leda Cosmides and John Tooby

We appreciated Ron Mallon's very thoughtful and interesting response to our paper and now worry from the tone of his response that he interpreted it as being disrespectful to the project of deontic logicians. Our hope instead had been to see whether evolutionary approaches could potentially add something to their work, by bringing some pertinent results to their attention and exploring how the deontic logic enterprise might proceed in light of them.

Dueling Oughts

Mallon raises the possibility of creating a domain-general deontic logic that does not entail domain-specific inferences but that does not violate them either. We had assumed that at present the project of deontic logicians was to create an inferential system that is completely general and sufficient to generate all deontic judgments. In arguing that a domain-general deontic logic would violate inferences made by social contract algorithms or a precautionary system, what we had in mind was a deontic logic sufficient to encompass these domains and make all the inferences that social contract algorithms and precautionary mechanisms do. If, instead, the goal is modified to produce a more limited domain-general logic that complements the operation of these domain-specific systems, then Mallon may be right in arguing that this is possible.

However, we are still not clear how this more limited and complementary deontic logic would resolve conflicts that arise between the injunctions of functionally distinct "oughts," such as OUGHT_{SC} and OUGHT_{PREC}. Let's consider Matviyko's smoking dilemma, posed by Mallon. Promises need not be social exchanges, but let's say Matviyko's agreement with his wife was one, something like, "If I (wife) support you this year, then you must finish your novel by the end of the year." As Mallon points out,

W



people resolve dilemmas like Matviyko's every day. But are these dilemmas usually resolved by deontic reasoning? Could they, in principle, be resolved by a deontic logic? Let's consider Matviyko's options.

1. Technical fixes often work: Matviyko could satisfy both oughts with a supply of nicotine gum.
2. Matviyko's wife, deciding she cares more about his health than his promise, could release him from his obligation to finish by the end of the year. This is part of the logic of social exchange (section IV.vi.ii).
3. Matviyko could prioritize his *values*: which is more important to him, keeping his word to his wife or beginning immediately the process of improving his health? Or his decision could be regulated by computations that engage a welfare trade-off ratio (Tooby & Cosmides, 2005), a variable specifying how much of his own welfare he is willing to trade off for that of his wife.

These value judgments could be made without engaging a moral reasoning system or consulting moral principles (and evolutionary psychology has a great deal to say about how welfare trade-off ratios should regulate decisions). Alternatively, moral principles could inform and shape these value judgments: values, utilities, and value hierarchies can surely be part of moral reasoning in general. But are they part of deontic logic? How would $\text{OUGHT}_{\text{General}}$ resolve a conflict between OUGHT_{SC} and $\text{OUGHT}_{\text{PREC}}$ *without* reference to values? And if values are necessary, perhaps something more than deontic logic is needed. Subjective expected utility theory integrated a calculus of utilities with the probability calculus. Perhaps moral philosophers are already working on a similar integration of values with a deontic calculus.

If not, here is another way deontic logicians may profit from considering certain aspects of evolutionary psychology. We have proposed that the human mind contains a computational system that infers and evaluates welfare trade-offs (Tooby & Cosmides, in press-b). It defines the conditions that elicit anger, guilt, affection, and other emotional/motivational systems, and interprets social interactions in terms of the welfare trade-off ratios implied by each person's actions. This raises the possibility that deontic reasoning is produced not by deontic procedures operating in isolation, but by deontic procedures supplemented by outputs from the welfare trade-off interpretive system. That is, having agreed to pick up a visiting speaker at the airport, I am obligated to do so—but not at the cost of dying. There are an infinite number of side constraints on every obligation, and these might be economically captured by a system of a few

welfare trade-off variables operating in an implicit motivational/interpretive system.

Library Model of Cognition

Referring to Richard Samuels's library model, Mallon points out that domain-specific inferential results need not imply domain-specific inference procedures (which is true). He then suggests that the domain-specific pattern of results for social exchange could be produced by domain-general computational procedures operating on a library of evolved, domain-specific information about social exchange: "For example, we might posit that ordinary people have an innate 'library' of *mentally represented rules*, such as, IF PERSON A SETS A CONDITION C ON THE BORROWING OF A'S PROPERTY P, THEN IF PERSON B SATISFIES C, THEN A OUGHT NOT TO MIND THAT B BORROWS P" (p. 124, italics added).

But is this actually different from what we have proposed? Here are three possible interpretations.

1. *Mentally represented rules* Mallon refers to this library entry as having "mentally represented rules," and their content is clearly about social exchange—*B's borrowing X* implies that *X* is a *benefit* to borrower *B*, so the rule has to do with what *A* requires in exchange for providing a benefit to *B*. Thus, Mallon has proposed a library of "mentally represented rules" dealing with social exchange; social contract theory proposes a set of such rules. If what is in Mallon's library are rules of inference, then his proposal is just a different way of making the same claim that social contract theory does. After all, at some stage of processing (audition, word recognition, etc.), some relatively domain-general computational processes must interact with social contract algorithms. Our claim is just that those processes cannot, by themselves, account for the rules of transformation that people apply in interpreting and reasoning about social exchange. Domain-specialized rules are needed.

2. *Ought types and inference* Perhaps Mallon is proposing a library with a *different* set of domain-specific rules than social contract theory proposes. For example, "Then *A* ought not to mind that *B* borrows *P*" tells us nothing about what *A* will mind, that is, it tells us nothing about what *A* will view as cheating. This means the rule given does not capture the reasoning patterns we have found (cheater detection; relaxing cheater detection when condition *C* was not satisfied by accident, etc.).

But let's leave that empirical matter aside: presumably Mallon thinks the library contains more mentally represented rules than this. But he then

proposes that the token representation of *ought* appearing in this rule may be type-identical to an ought-token appearing in inferential rules for precautions or other domains. This, he argues, would allow for the possibility that the brain contains only one type-representation of “ought,” which then acquires “different inferential roles in social exchange and precautionary domains because of the presence of ‘libraries’ of mental representations encoding information about those domains” (p. 124).

Mallon does not seem to object when we identify the meaning of a concept with the inferential role it plays, so on his account, the type-representation of “ought” would activate a set of domain-independent inferences and then trigger *extra* inferences depending on the inferential role it plays within each library entry in which it appears. Whether there exists a set of domain-independent *ought* inferences that are capable of meshing with these libraries is an empirical question, one which deontic logicians are well-placed to answer. But even on this account, OUGHT_{SC} and OUGHT_{PREC} still exist. These concepts are, precisely, the different bundles of ought-related inferences that are triggered when a library entry containing the type-representation of *ought* is activated. To be a sufficient account of deontic reasoning, the entire bundle is relevant, and each bundle should be different, even if a subset of inferences appear in each bundle.

3. *Declarative knowledge in the library* Rather than containing rules of inference, the library entry on social exchange might be a set of inert facts, a packet of declarative knowledge that is “looked up” by domain-general procedures. (This interpretation is most consistent with the library metaphor.) Unfortunately, this version of the library model has an empirical problem: it does not explain why performance elicited by social contracts and precautions is so much higher than that elicited by other deontic rules or, indeed, by familiar indicative conditionals.

The selection task asks subjects to *search* for potential violations of a conditional. For domain-general procedures to search for violations, they must be able to look up what conditions count as a violation. And people have this knowledge, even for indicative conditionals: they have a library entry specifying that the combination of *P* and *not-Q* violates indicative rules (see section I.ii). Yet they do not spontaneously search for information that could reveal potential violations of indicative conditionals, even when they are explicitly asked to do so. In fact, they perform poorly even when we call their attention to what counts as a violation (see section I.ii). This means that *searching* for violations requires more than domain-general procedures plus a library entry specifying what counts as a violation.

It follows that having a library entry specifying what counts as cheating is not sufficient to make people good at searching for cheaters. Computationally, something more is needed. People know what counts as a violation for rules drawn from many domains, but this knowledge does not explain when they succeed and fail in their search for violations. This failure of library models was a major impetus for proposing a cheater detection mechanism, equipped with procedures that direct information search in a way likely to reveal cheaters.

Deontic Specialists (or Why Care about the Epidemiology of Representations?)

We agree with Mallon: specialists, including deontic logicians, may be pursuing valuable enterprises even when these conflict with popular belief—as Mallon points out, scientists do not abandon Darwinism for creationism just because evolved inference systems have trouble with deep time. But there is a difference between moral claims and truth claims. Truth claims, however esoteric and nonintuitive, imply facts that can be validated by empirical investigations. But what validates the truth of a moral claim? Deontic logicians do not need to answer Hume’s penetrating question if their goal is to construct a deontic logic that captures the moral intuitions of a species or culture. But if they eschew this goal, concentrating instead on constructing various esoteric formalisms—ones whose conclusions conflict with one another (as well as with folk intuition)—then what? What criteria will they use to decide which of these mechanical formalisms is the right one, the one that objectively encapsulates the logic of moral obligation? Without an answer to this question, it is not clear how that particular enterprise can succeed.

There is another reason to consider the mesh between the population intended to profit from deontic discoveries and the esoteric formalisms that deontic specialists produce. Mallon argues that the public at large need not understand what specialists are thinking for their intellectual products to influence our lives, citing scientific discoveries as an example. True. But the paths by which scientific and deontic discoveries influence our lives differ in ways that are causally and morally important. Causally, I don’t need to know how my car engine works (because engineers have indeed considered how to make scientific discoveries mesh with my intuitions). But I do need to understand a moral reasoning system to apply it to my decisions. The alternative is a set of deontic experts (or a computer program) making moral decisions for the rest of us. Deontic experts influence us

with persuasion and sometimes through coercion (using the apparatus of the state). But in either case, the epidemiology of representations remains relevant. What happens when political and legal systems run by specialists systematically produce moral judgments or social outcomes that deeply violate evolved moral intuitions? It does not seem far-fetched to suggest that they eventually will lose public support.

So yes, deontic logicians are free to ignore the design of evolved deontic reasoning systems, but we still believe they are ill-advised to do so.