

Further evidence that normative versus non-normative responses tell us little about the mechanisms that produced them comes from a series of studies recently conducted in my lab. These studies employ a paradigm developed to study processes that monitor inferences and determine when additional analytic thinking is warranted. In this paradigm, participants are asked to give an initial, intuitive answer to a problem, make a metacognitive judgement about how right that answer feels, and are then allowed to take as much time as needed to give a final answer. The metacognitive experience that accompanies an initial decision predicts a lot about the effort that people put into solving a problem. Specifically, a strong feeling of rightness about a decision determines the amount of time spent thinking about a problem and whether or not the initial answer is changed in favour of another. Importantly, a strong initial feeling of rightness does not reliably predict whether the final answer is normatively accurate, despite the fact that it exerts substantial control over subsequent analytic reasoning. Moreover, in several studies, we have observed that when people do change their answers, they are often just as likely to change from a normatively correct answer to a wrong one, as vice versa (Shynkaruk & Thompson 2006; Thompson et al., under review). Understanding what motivates an answer change, what constraints an answer must satisfy to be retained, the information that is recruited to rework the answer, and so on, will tell us a lot about human reasoning. Knowing that the final outcome is normative tells us virtually nothing about the underlying mechanisms.

Neurath’s ship: The constitutive relation between normative and descriptive theories of rationality

doi:10.1017/S0140525X11000586

Michael R. Waldmann

Department of Psychology, University of Göttingen, 37073 Göttingen, Germany.

michael.waldmann@bio.uni-goettingen.de

<http://www.psych.uni-goettingen.de/waldmann>

Abstract: I defend the claim that in psychological theories concerned with theoretical or practical rationality there is a constitutive relation between normative and descriptive theories: Normative theories provide idealized descriptive accounts of rational agents. However, we need to resist the temptation to collapse descriptive theories with any specific normative theory. I show how a partial separation is possible.

The target article by Elqayam & Evans (E&E) reminds us of a plausible methodological norm: Do not confuse normative with descriptive theories! In many areas of psychology this distinction seems trivial. Models of memory or perception try to elucidate how these systems actually function, not how they should function. But what about theories of reasoning and acting, the classical domains of theoretical and practical rationality? Is it prudent to cleanly separate normative and descriptive theories here as well? My claim is that in psychological theories of rationality there is an intrinsically constitutive relation between normative and descriptive theories (Spohn 2002). I exemplify this claim in two different psychological research areas, one from the domain of theoretical, the other from practical rationality. I also argue that we need to resist the temptation to collapse descriptive theories with any specific normative theory.

Theories of *theoretical rationality* are concerned with how we achieve the goal of acquiring correct beliefs about the world. The close ties between normative and descriptive theories can easily be seen when we consider a scientist conducting an experiment. To explain the observed behavior, for example, randomization of subjects, it seems natural to invoke normative theories of how

experiments *should* be conducted. The reason why normative theories work as descriptive explanations here is that we understand the scientist as a rational, goal-oriented person. Everyday reasoning is, in contrast, rarely guided by explicit methodological knowledge. Nevertheless, people’s beliefs also have a normative force. People view themselves as motivated by reasons, and strive for optimality to achieve their goals. They distinguish between true and false beliefs; knowing they may err, they revise their beliefs and accept corrections. We would not say that a person believes that *A causes B* without simultaneously assuming that this person considers this proposition to be true. In sum, people conceive of themselves as rational agents, which make normative theories the natural candidate to explain their behavior.

Recent research about *causal reasoning* has obtained a wealth of evidence showing that we try to go beyond observations to obtain knowledge about causal relations in the world. People are rarely aware of how they acquire causal knowledge, but they understand what it means to respond to a causal query. Thus, a natural place to look for candidate explanations is normative theories of causal inference. In fact, virtually all currently competing theories of causal reasoning can be structured according to the preferred normative account motivating the theory (Waldmann & Hagmayer, in press). Causal reasoning cannot be modeled without some normative theory that tells us what causal judgments are.

Moral reasoning, which belongs to the domain of *practical rationality*, is another example of the constitutive relation between normative and descriptive theories. When we request moral judgments from subjects we do not want to learn about their preferences, wants, or inhibitions; rather, we want them to provide a normative evaluation of whether an act is *right* or *wrong*. Again, we conceive of subjects as rational agents who offer us responses to normative requests. Therefore, it is not surprising that an overview of research on moral judgment reveals that psychological theories use concepts from normative theories of morality (Waldmann et al., in press). For example, when interpreting responses to the famous trolley problem about whether it is permissible to sacrifice one person to save five, many theories focus on acts, outcomes, or values, concepts that have been highlighted as morally relevant in normative theories. And even if other non-moral factors are included in the explanations, we still interpret subjects’ responses as driven by the motive to provide a normative assessment. Otherwise, we could not say anymore that we are studying moral judgments.

Although I claim that in theories of rationality, normative theories are constitutive for the theoretical description of the target phenomena, I agree with E&E that there is a danger of oversteering the empirical adequacy of specific theories. Normative theories come in many variants, which compete. Their development is governed by factors such as coherence and consistency, which are less influential in everyday reasoning. Moreover, normative theories may restrict themselves to aspects of the target domain that are less relevant in everyday reasoning. Hence, it is unlikely that any specific normative position can be directly used as a descriptive theory. Responses in causal reasoning tasks have been interpreted as evidence for Bayesian causal network models, although it may often be possible to provide a more parsimonious account for individual phenomena by stripping away unnecessary untested assumptions implicit in these models. Similarly, the interpretation of responses as consequentialist or deontological in the trolley problem can often be more parsimoniously explained without attributing these global philosophical positions to subjects.

A sensible research strategy, therefore, seems to be to use one of the competing normative theories as a starting point, but to then ask whether all the assumptions inherent in these theories are empirically validated and necessary to explain the target behavior. Later, we can even go outside the realm of rationally relevant explanatory concepts. Rationality is certainly an idealization of thinking; many of our thoughts and actions are

influenced by factors that, on reflection, nobody would consider as legitimate. For example, Eskine et al. (2011) found that moral judgments about transgressions (e.g., stealing; taking a bribe) tend to be harsher when subjects were drinking a shot of a bitter beverage than when they were given water or a sweet beverage. The taste of a beverage certainly does not constitute an acceptable argument for a moral judgment.

In short, it may be necessary to rebuild the selected normative framework, delete components, sacrifice coherence and consistency, and even add non-normative factors. Some core components will stay invariant, to guarantee that we still model the target competency, such as causal or moral reasoning. The end result of the revision process may be a caricature of any accepted normative theory, but its normative foundation will still be discernible. Or as the philosopher Otto Neurath famously claimed: “We are like sailors who have to rebuild their ship on the open sea, without ever being able to dismount it in dry-dock and reconstruct it from the best components” (see Quine 1960, p. vii).

What is evaluative normativity, that we (maybe) should avoid it?

doi:10.1017/S0140525X11000604

Jonathan M. Weinberg

Department of Philosophy, University of Arizona, Tucson, AZ 85721-0027.
jmwainberg@email.arizona.edu

Abstract: Elqayam & Evans (E&E) argue that we should avoid evaluative normativity in our psychological theorizing. But there are two crucial issues lacking clarity in their presentation of evaluative normativity. One of them can be resolved through disambiguation, but the other points to a deeper problem: Evaluative normativity is too tightly-woven in our theorizing to be easily disentangled and discarded.

Elqayam & Evans (E&E) advise us to jettison the normative element in theoretical psychological investigations. To follow their advice, though, we need to know just what it is we are to jettison. There are two crucial areas of unclarity in their presentation of normativism. The first can simply be disambiguated, but the second indicates a deeper problem with their project.

First, some of their formulations of normativism involve a formal system (as in their Abstract: “human thinking reflects a normative system against which it should be measured and judged”). Elsewhere in the target article, normativism involves an appeal to something that is “a priori,” “unconditional,” or “universal.” But at other points, systems drop out, and all that is involved is a particular kind of “oughtness”; and indeed, at other points in the text, merely an appeal to rightness and wrongness, or “error” is deemed sufficient to count as normativist. It is dangerous to elide the difference between the notion of formal or a priori systems, and evaluative normativity itself, for there are evaluative normative theories that traffic in oughtness, but not on the basis of any such formal or a priori principles; epistemic reliabilism, for example (Goldman 1979; Kornblith 2003), and, as I will argue in a moment, the ecological rationality of Gigerenzer, Todd, and others. Some disambiguations are therefore in order.

E&E’s arguments mostly aim at universality (e.g., the “arbitration problem”) and the is-ought/ought-is inferences. So let’s tease a apart several different sub-claims of their anti-normativism that might be at issue:

(*Humeanism*) We should not make any hasty inferences from “is” to “ought,” or vice versa.

(*Anti-Universalism*) There is no unique correct evaluative normative framework applicable to all persons and situations.

(*Descriptivism*) We should not make any substantive (as opposed to merely “inspirational”) inclusion of evaluative normativity as a part of our psychological theorizing.

The authors’ arguments for the first two anti-normativisms are generally well-taken (see also, Weinberg 2007). But in some places they are perhaps over-zealous. For example, there are surely other explanations available for the focus on naive as opposed to trained subjects, on the very plausible presupposition that only a small portion of the population receives any such training; and so what we learn from studying such subjects may not generalize well. And, for that matter, expert reasoning populations have long been studied as well, as with Herbert Simon’s highly influential research. Also, the authors’ discussion of Humeanism could perhaps be more sensitive to a minor quandary they put themselves in: Having rejected a priorist approaches to the question of evaluative normativity, they either have to endorse some “is” facts as evidence for “ought” facts, or end up as full-blown skeptics about such normativity.

Disentangling Descriptivism from Humeanism and Anti-Universalism, however, does not yet render it clear, for “evaluative normativity” itself needs further clarification. One worry is that E&E have not actually succeeded in identifying a distinct form of normativity that can be cleanly set aside. They distinguish it initially from instrumental, bounded, ecological, and evolutionary forms of normativity, which they find unproblematically descriptive in nature. But they are too hasty in assimilating ecological normativity, for psychologists in that school seem to traffic in forms of correctness that cannot be boiled down to “Darwinian and Skinnerian algorithms.” Ecologists’ writings are rife with discussions of a “fit” between environment and mind, and of “success” or simply “good reasoning” that is neither constituted by (even if perhaps highly correlated with) the satisfaction of specific desires of individuals nor grounded in any appeal to what has historically promoted, or would today promote, reproductive fitness. One way of seeing how ecological normativity cannot be explicated in terms of instrumental normativity, is that the former is very often analyzed in terms of the relationship between a heuristic and an environment, without reference to any (even hypothetical) desires of some agent (see Over 2000). So, although E&E place ecologists in the “low normativism” zone, this is due to the ecologists’ conflating evaluative normativity itself with the use of normative systems; ecological rationality eschews the latter, but is robustly committed to the former. This also explains the difficulty E&E have with those researchers apparently deploying “oughtness” more than their location in the “low normativism” zone would expect.

E&E also go on later to invoke yet another form of allegedly unproblematic normativity, “epistemic rationality, in the sense of holding well-calibrated beliefs” (sect. 5.2, para. 4). But such calibration is not an unevaluative notion. Indeed, the logicians or Bayesians could easily claim that such normativity is exactly what they themselves are theorizing in terms of. For example, the reason, according to logicians, why you should not affirm the consequent is that it will generally lead you to believe falsehoods. The reason, according to the likes of Kahneman and Tversky, why heuristics tend to lead to biases is that they leave us susceptible to various forms of false beliefs.

There is a fundamental problem here: The sort of normativity that E&E want to set aside is simply too tightly wound around our psychological theorizing to be disentangled and discarded. That the authors find themselves writing in terms of instrumental rationality where agents and goals drop out of consideration altogether, and of the same epistemic rationality that their opponents would embrace, suggests that they are just as deeply enmeshed in evaluative normativity as everyone else, if in an unintentionally cryptic way. The authors are concerned that normative thinking has a too powerful, “biasing” influence on scientists’ minds. That very fact, however, suggests that trying to get scientists to swear off of normative thinking altogether may be simply impossible. Trying to do without it may only result in