

“Error Theory”
Entry for *International Encyclopedia of Ethics* (Wiley-Blackwell, forthcoming)
Richard Joyce

To hold an error theory about morality is to endorse a kind of radical moral skepticism—a skepticism analogous to atheism in the religious domain. The atheist thinks that religious utterances, such as “God loves you,” really are truth-evaluable assertions (as opposed to being veiled commands or expressions of hope, etc.), but that the world just doesn’t contain the items (e.g., God) necessary to render such assertions true. Similarly, the moral error theorist maintains that moral judgments are truth-evaluable assertions (thus contrasting with the noncognitivist), but that the world doesn’t contain the properties (e.g., moral goodness, evil, moral obligation) needed to render moral judgments true. In other words, moral discourse aims at the truth but systematically fails to secure it. If there is no such property as *moral wrongness*, for example, then no judgment of the form “X is morally wrong” will be true (where “X” denotes an actual action or state of affairs). Advocates of this position include Hinckfuss 1987; Joyce 2001; Mackie 1977 (see MACKIE, J. L.). Various forms of moral skepticism—some of which are arguably instances of the error theoretic stance—have been familiar to philosophers since ancient times. (See SKEPTICISM, MORAL.)

Error theoretic views can be controversial—as in the case of religion and morality—or widely agreed upon—as in the case of ghosts and phlogiston. It is important to note that error theorists maintain that the judgments in question are erroneous not merely because of the absence of any objective moral facts sufficient to render them true, but also because of the absence of any *non-objective* moral facts sufficient to render them true. There is, for example, a kind of moral realist who maintains that moral properties are objective features of the universe (see REALISM, MORAL). There is also a family of metaethical views according to which moral properties are in some manner constituted *by us*—by our beliefs, attitudes, practices, etc. We might call the latter view “constructivism” (see CONSTRUCTIVISM, MORAL). The error theorist opposes both positions, resisting the proposal that moral wrongness (say) is a non-objective property as resolutely as he or she denies that moral wrongness is an existing objective property.

Why would someone maintain such a view? Isn’t holding a moral error theory a pernicious and appalling position to adopt? Are we going to take seriously philosophers who maintain that there was nothing morally wrong with Hitler’s extermination of the Jews?!

These common and understandable reactions must be treated with caution. Consider the last claim: that moral error theorists don’t believe that there’s anything morally wrong with genocide. Strictly speaking, this is correct. On the other hand, this should not be acknowledged without also noting that the moral error theorist is just as keen to deny that genocide is morally good or morally permissible as he is to deny that it is morally wrong. The error theorist denies that anything is morally *anything*. In the context of everyday conversation, to deny that there is anything morally wrong with X implies that one thinks that X is morally good or morally permissible—but a conversational context in which radical moral skepticism is live contender is precisely one where that implication breaks down. Attempts to embarrass the error theorist by forcing him to admit that he doesn’t think that there is anything wrong with [insert your favorite crime here] trade on equivocating between these contexts.

Is a moral error theory an appalling and pernicious position to adopt? In so far as most people are adamantly *not* error theorists, then we must accept that it is apt to appall. But whether adopting and espousing an error theory is pernicious is really an empirical question that has not been tested. The general concern is, presumably, that various socially desirable behaviors that receive widespread moral endorsement will be less likely to be performed if

that moral backing is removed. And, *mutatis mutandis*, socially undesirable behaviors will be more likely to be performed. There may or may not be a sound basis to this fear; it is an empirical question. An important consideration here is the fact that there are many effective ways of encouraging or discouraging action that appear to have nothing to do with morality. A person may greatly *desire* to live in a world without genocide (say), and it is not obvious that the force of that desire must be diminished if the person resists categorizing genocide in *moral* terms. Compare a person's desire to protect his or her children from harm. There is no *a priori* reason for assuming that "non-moralized" preferences must be weak or whimsical. By analogy (again with atheism): There no doubt was (and in some circles still is) a fear that a widespread rejection of theism would lead to general social decline. But this does not seem borne out by data; it is commonly allowed that a purely secular morality can undergird social cohesion in a perfectly effective manner. When addressing the fear that a widespread rejection of morality would lead to a general social decline we should be open to the possibility that the same pattern might hold: that a purely non-moral suite of entrenched attitudes, preferences, and habits, etc., could step in and do the same job just as effectively, or perhaps even better. (See Garner 2009; Hinckfuss 1987.)

Just as the adoption of an error theoretic stance towards morality has no logical implications for individuals' behavior or preferences, nor does it have any logical implications regarding what speakers *do* with their faulty moral vocabulary. The natural assumption is that the error theorist should cease to employ moral discourse—a position known as "eliminativism." Of course, even the eliminativist will continue to utter moral words on occasion—in making such claims as "There are no such things as moral rights" or "Bush believes in evil"—but he won't employ moral sentences that imply the existence of moral properties (such as "X has moral rights" or "Y is evil"). However, the error theorist is not forced by any logical consideration to embrace eliminativism. Perhaps epistemological norms will proscribe the continued *assertion* of such moral sentences (assuming that a condemnation of lying is an epistemological norm), but that is not to say that they proscribe their mere utterance. Instead of being an eliminativist, the moral error theorist may adopt a *fictionalist* attitude. (See Joyce 2001; Kalderon 2005.) The fictionalist carries on using much or all of moral language, all the while believing morality to be an erroneous system. (This, at least, is true of the so-called revolutionary fictionalist. The hermeneutic fictionalist, by contrast, thinks that our attitude towards morality has always been akin to pretense and hence is no error theorist.) The key to pulling off the trick of carrying on with morality without believing it, without contradiction and without violating any epistemological norms, is to adopt a kind of acceptance that is not belief and to employ a kind of speech act that is not assertion. Since such categories exist, error theoretic fictionalism about morality is logically coherent, but whether it is a psychologically viable option, and, if so, whether it would be a pragmatically good idea, are further unresolved empirical issues.

Having noted that adopting a moral error theory may not have the detrimental practical effects as is often assumed at first glance, we now turn to what positive considerations might lead one to endorse it. It seems fair to say that the usual motivation for the moral error theory is dissatisfaction with the alternatives, which here are classified as noncognitivism, constructivism, and realism. Since there are many criticisms that may be leveled at each of the alternatives, the following discussion is intended to give a flavor of the error theorist's thinking, rather than present any persuasive arguments.

The noncognitivist maintains that moral sentences should not be interpreted as aiming at the truth in the first place. (See NONCOGNITIVISM.) One kind of noncognitivist (who might better be called a "nonfactualist") holds that although moral sentences typically have a propositional grammatical structure (e.g., "Stealing is wrong") in fact their real logical status is non-propositional; they can be translated into something else (e.g., "Boo to stealing" or

“Don’t steal”) for which the question of *truth* shouldn’t arise. (See e.g., Ayer [1936] 1971; Blackburn 1998.) (See EMOTIVISM; PRESCRIPTIVISM.) Another kind of noncognitivist allows that moral sentences really are propositional but maintains that we don’t *use* them to make assertions; we use them to perform some other kind of speech act, such as expressing our approval, or commanding compliance, or revealing our adherence to a normative framework. (See Kalderon 2005 for this distinction.)

The error theorist finds noncognitivism unconvincing. There is, however, something that the error theorist and the noncognitivist share: a conviction that if we take moral predicates at face value, then the enterprise of pinpointing actual moral properties (badness, obligatoriness, permissibility, etc.) to assign to those predicates—an enterprise necessary if our moral judgments are to be true—is doomed to fail. The noncognitivist’s response is that we should not take moral predicates (or the speech acts that employ them) at face value, and thus the perennial search for moral properties may be called off. The error theorist’s response is that there are no actual properties to be assigned to the predicates, and thus moral truth is unavailable.

What makes many philosophers (including many error theorists) uneasy about the noncognitivist’s response is that it smacks of interpreting a discourse in an eccentric manner simply to avoid philosophical difficulties. If there were some independent evidence that moral predicates operate in an unusual way, some concrete evidence that moral utterances are something other than common-or-garden assertions, then the noncognitivist’s case would be greatly strengthened.

In addition, many philosophers (including many error theorists) find that noncognitivism fails to adequately satisfy certain metaethical desiderata. It has trouble accounting for the authority of morality: If S’s utterance of “Stealing is morally forbidden” amounts to no more than an expression of S’s feelings (“Boo to stealing!”) then why should anyone who is not antecedently inclined to care about S’s feelings pay any attention? If the audience responds “Hurray for stealing!” then any further discussion of the matter seems pointless. (That, admittedly, is a rather crude version of noncognitivism, but it is not obvious that more sophisticated versions do much better.) There is also a longstanding concern about whether the noncognitivist can account for certain basic semantic features of moral discourse. Most famous is the worry that if the underlying grammar of (say) “Stealing is wrong” is not “X is P” but rather something like “Boo to stealing!” then how is the sentence going to operate in logically complex contexts, such as the antecedent of a conditional? For instance, the sequence of words “If boo to stealing then boo to stealing from John” seems puzzling, to say the least; it does not even qualify as a sentence. The orthodox truth-functional rules that allow us to understand what an “if...then...” statement mean do not seem to apply. (See FREGE-GEACH OBJECTION.)

If noncognitivism is found inadequate (for whatever reasons), then we are left with two broad alternatives to the error theory: realism and constructivism. (See REALISM, MORAL; CONSTRUCTIVISM, MORAL.) Both views take moral predicates and speech acts at face value, and both think that moral assertions will often be true. Where they diverge concerns what they think is the nature of those properties that render these assertions true. Roughly, realists think that true moral assertions are true in virtue of the instantiation of certain objective properties, whereas constructivists think that we (our attitudes, practices, conventions, etc.) have an ineliminable and important role in the making of moral properties. Speaking even more roughly, for realists morality is discovered, for constructivists it is invented. Many philosophers are pessimistic that this distinction can really withstand critical scrutiny (see Rosen 1994), and some constructivists consider themselves realists, but let us here sideline these complications and outline what the error theorist is likely to find objectionable about each option.

The error theorist's dissatisfaction with constructivism is likely to mirror one of the sources of dissatisfaction with noncognitivism—namely, that it seems to leave the question of why anyone should care about morality unsatisfactorily open. Consider a crude version of constructivism for illustrative purposes: Suppose that “X is morally wrong” is given the relativistic interpretation “I disapprove of X.” Hearing another agent report this fact about her feelings, why should anyone care? Why should anyone (even someone who approves of X) feel inclined to debate the matter? The mere fact that people feel approval and disapproval towards certain things doesn't seem to constitute or undergird the kind of binding normativity with which we seem ordinarily to imbue morality.

Constructivism can be made less crude along several dimensions. Instead of the idiosyncratic “I,” reference might be made to the collective response of a group of people. These individuals might be improved in various ways: provided with full information, ambitions to live in harmony, increased powers of reflection, impartiality, and so forth. The relevant response may be something more subtle than the blunt “disapproval.” And all this idealization necessitates that the modal relation be altered: Instead of X being morally good if such-and-such folks *do* have a certain kind of response to X, it will depend on whether these folks *would* have that response, were they to be around to witness or contemplate X. (See IDEAL OBSERVER THEORIES; RESPONSE DEPENDENT THEORIES.)

The error theorist remains doubtful, however, that these various “improvements” address the concern. Suppose Mary recognizes that a certain action, ϕ , is such that fully informed persons who share the aim of reaching free and unforced agreement would endorse a norm in favor of ϕ . The presence of that dispositional property *may* be of interest to Mary, but it may not be. Certainly its presence doesn't seem to constitute any kind of practical *demand*. Mary may be as unmoved by its instantiation as she is when she reflects that ϕ is also such that drunken Vikings would mock its performance. There seems nothing in the presence of the former disposition per se that *demand*s Mary's attention and compliance. Even if Mary *wants* to be a fully informed person engaged in reaching agreement with other fully informed agents (in a way that she doesn't want to be a drunken Viking), the presence of the dispositional property does not demand her compliance. (As I sit in my office I may yearn to be lying on a beach, but that doesn't give me reason to put on sun block now at my desk.) To recognize this is to acknowledge that, according to the constructivist, moral properties—rightness and wrongness, virtue and duty—don't really constitute demands; they are properties that an agent need take an interest in only if she is antecedently inclined to care about certain things. This lack of practical “oomph” in constructivist properties is likely to be what leaves the error theorist unconvinced. (Of course, it may be something else entirely that leaves him unimpressed.) The error theorist is likely to maintain that some kind of special normative “oomph” is a distinctive and essential part of our very moral concepts, such that the anthropocentric dispositional properties offered by the constructivist simply cannot *be* the constituents of the moral realm. (See Joyce 2001: chapter 2; Mackie 1977: 27-35.)

In short, the practical relevance of any constructivist property comes from internal sources—all versions of constructivism seem to make the normative importance of moral rightness and wrongness depend on our prior attitudes—and the error theorist judges that this fails to accommodate a key desideratum: that moral norms and values are imbued with a kind of force that demands compliance *regardless of whether we care*—demands that come from *without*. By contrast, placing moral properties in an *external* realm, allowing them to be objective constituents of the world, is the distinctive move of the moral *realist*. The error theorist applauds that moral realist for seeing this—acknowledging that the realist has a better appreciation than the constructivist of the conceptual commitments of moral discourse. In fact, the error theorist may think that the moral realist understands morality perfectly well *at*

a conceptual level. But whereas the moral realist thinks that there are properties in the world that actually answer to these conceptual commitments, the error theorist thinks otherwise.

Compare our concept of the Loch Ness monster. It is the concept of an objectively existing thing. Nobody maintains that the Loch Ness monster has the unusual ontological status of being such that if a sufficient number of people believed in its existence then it would somehow pop into being, as a “subjective” entity (like the value of gold, say). Realists and skeptics alike are adamant that any constructivist interpretation of the Loch Ness monster’s status can be rejected (as can be any noncognitivist interpretation of Loch Ness monster discourse). Where the realist and the skeptic diverge is the all-important matter of whether there actually is something swimming around that answers to the objectivist concept.

Sticking for the moment with an imperfect analogy: Why would one be skeptical of the existence of the Loch Ness monster? A reasonable answer would be something like: Because systematic attempts to find positive evidence have largely failed, because hypotheses explaining how there could exist such an odd creature are implausible given what else we know about the world, and because what positive evidence has come forward is easily explained away. This very roughly reflects the error theorist’s attitude towards moral realism. He deems that the numerous attempts to delineate a set of objectively existing moral properties have all failed, that the idea that there exist such properties is implausible (that such properties would be very odd), and that there is no evidence that cannot be easily explained by alternative “non-moral” hypotheses. (See QUEERNESS, ARGUMENT FROM.)

Note how the error theorist’s standard argument has two steps. First there is a conceptual step of establishing what moral properties would have to be like. We have seen that a reference to “objectivity” (or some equivalent notion) is likely to appear in the error theorist’s conceptual analysis, thus excluding the constructivist from the running. (Of course, other things will appear in the analysis too; if moral properties were *just* objective, then the realist could simply point to cats or tables.) The second step is ontological: establishing that there do not exist any properties answering to the concepts in question. The second step might be established through *a priori* methods (showing the concepts in question to be incoherent, for example) or through *a posteriori* methods (concluding that the requisite kinds of properties do not fit with an accepted worldview attained through empirical effort). The systematic failure that the error theorist attributes to moral discourse might be either a necessary failure (such that the error theory holds at every possible world) or a contingent affair. For example, a moral error theorist who is also an atheist might concede (A) that the existence of an all-powerful, omni-benevolent, etc., deity would suffice to underwrite the existence of moral properties, and (B) that atheism is only contingently true.

The error theorist is not obliged to employ this type of two-step argument in favor of the position. She might, for example, instead be “pushed” into accepting an error theory entirely through becoming thoroughly disillusioned with all alternatives, in which case she need endorse no specific thesis about what moral properties conceptually would have to be like (except, perhaps, denying that moral concepts are concepts of items that exist if a sufficient number of people believe that they exist). Nevertheless, it seems fair to say that the two-step argument is the error theorist’s standard strategy—to such an extent that several critics have had trouble distinguishing what it takes to be an error theorist from what it takes to endorse this kind of argument (see Dreier 2005; Finlay 2008). The two steps of the error theorist’s standard argument throw up two kinds of opponent. One opponent tackles the error theorist head-on: arguing that the moral concepts outlined by the error theorist are in fact defensible (e.g., arguing that the world really does contain objective norms that bind our actions irrespective of our attitudes). The other opponent concedes that such properties really are unacceptably odd and that we should not believe in them—but insists that it is a mistake to

think that morality was ever essentially committed to such oddities. Even if many people have believed in this mysterious “binding force” (the concessive opponent maintains), this would not suffice to make it an essential ineliminable component of our moral conceptual framework. The head-on opponent accuses the moral error theorist of making an ontological mistake: of underestimating what categories of things the world contains. The concessive opponent accuses the moral error theorist of a conceptual mistake: of overestimating what features are essential to moral concepts. We may, thus, sum up the standard moral error theoretic strategy by saying that it combines a rich view of moral concepts with a meager ontological view of the world. Some will complain that its conceptual theses are *too* rich, and many will protest that its ontology is *too* meager. But there is no obvious *a priori* or methodological pressure to take the same attitude to both domains (i.e., both rich or both meager), and there is no philosophical consensus on how either kind of dispute should be settled, and thus this form of radical moral skepticism remains a perennial contender.

SEE ALSO:

Constructivism, Moral; Emotivism; Frege-Geach Objection; Ideal Observer Theories; Mackie, J. L.; Noncognitivism; Prescriptivism; Queerness, Argument from; Realism, Moral; Response Dependent Theories; Skepticism, Moral

REFERENCES:

- Ayer, A. J. [1936] 1971. *Language, Truth and Logic*. Penguin.
- Blackburn, Simon 1998. *Ruling Passions*. Oxford University Press.
- Dreier, Jamie 2005. “Moral relativism and moral nihilism,” in D. Copp (ed.) *Handbook of Ethical Theory*. New York: Oxford University Press: 240-64.
- Finlay, Stephen 2008. “The error in the error theory.” *Australasian Journal of Philosophy* 86: 347-369.
- Garner, Richard 2009. “Abolishing morality,” in R. Joyce and S. Kirchin (eds.) *A World Without Values: Essays on John Mackie’s Moral Error Theory*. Springer Press.
- Joyce, Richard 2001. *The Myth of Morality*. Cambridge University Press.
- Hinckfuss, Ian 1987. “The moral society: Its structure and effects.” *Discussion Papers in Environmental Philosophy* 16. Philosophy Program (RSSS), Australian National University.
- Kalderon, Mark Eli 2005. *Moral Fictionalism*. Oxford: Clarendon Press.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. Penguin.
- Rosen, Gideon 1994. “Objectivity and modern idealism: What is the question?” in M. Michael and J. O’Leary-Hawthorne (eds.) *Philosophy In Mind*. Kluwer: 277-319.

SUGGESTED READINGS:

- Garner, Richard 1994. *Beyond Morality*. Temple University Press.
- Honderich, T. (ed.) 1985. *Objectivity and Morality*. Routledge & Kegan Paul.
- Joyce, Richard and Kirchin, Simon (eds.) 2009. *A World Without Values: Essays on John Mackie’s Moral Error Theory*. Springer Press.
- Liggins, David and Daly, Chris 2009. “In defence of error theory.” *Philosophical Studies*.
- Lillehammer, Hallvard 2004. “Moral error theory.” *Proceedings of the Aristotelian Society* 104: 93-109.
- Miller, Alexander 2003. *An Introduction to Contemporary Metaethics*. Polity.