

## LOGICAL LUCK

"Logic and ethics are fundamentally the same, they are not more than duty to oneself" (Otto Weininger). So goes the head quotation of Ray Monk's biography Ludwig Wittgenstein: The Duty of Genius. Monk thereby introduces Wittgenstein's peculiar admiration for the crackpot author of Sex and Character along with Wittgenstein's moralistic dedication to logic. Monk elaborates with anecdotes. For instance, Wittgenstein would pace Bertrand Russell's room mixing logic with self-criticism. Russell asked Wittgenstein whether he was thinking about logic or his sins. "Both!" barked Wittgenstein.

Wittgenstein's monism about the normative has a precedent; the rationalist tradition of ethics. Immanuel Kant said that lying was wrong because the attempt to universalize lying leads to a contradiction. He thought that the test of universalizability was a complete moral guide. Kant assumed that each agent can always detect a contradiction by introspection. Consequently, immorality is always a matter of incoherence, never a matter of ignorance. The moral status of the agent

depends solely on the properties he has on his own, never on the properties he has in virtue of relationships with other things.

If immorality entails irrationality and irrationality entails contradiction, then ethics and logic converge. I shall argue that this conditional convergence lurks behind the debate on moral luck and the debate about externalism in philosophy of mind.

### **1. Equivocation without guilt?**

We equivocate when we intend a term to be read univocally even though it is ambiguous. Deliberate equivocators are guilty of deception. More interestingly, unwitting equivocators also seem open to reproach. The discovery that one has equivocated prompts remorse for negligence or recklessness. The honest equivocator feels that the fallacy would not have been committed had greater care been exercised. The assumption is that one can always check whether one is using one's own terminology uniformly. In logic, one can be safe. Wittgenstein characterized religious experience as the feeling of being absolutely safe. (Monk 1990, 51)

Some philosophical positions imply that all equivocations are avoidable. The premises of Descartes's proof of the external world entail that God must always give each agent an opportunity to reason validly.

Descartes once remarked that intelligence must be one of the most evenly distributed of goods because no one complains for want of it. Descartes was not joking; he requires all error to be the product of freewill.

If equivocation is always avoidable, then we must always be able to detect synonymy in our reasoning. Michael Dummett draws an implication for the epistemology of language:

It is an undeniable feature of the notion of meaning -- as obscure as that notion is -- that meaning is transparent in the sense that, if someone attaches a meaning to each of two words, he must know whether these meanings are the same.

(Dummett 1978, 131)

This further implies that meaning depends on the intrinsic properties of the speaker's mind, that my meaning cannot directly depend on my environment or social situation. External conditions would be relevant only insofar as they are mediated by my narrow psychology.

However, semantic externalists have used twin earth thought experiments to argue that surprisingly many terms are indexical ('water') or are dependent on linguistic division of labor ('elm', 'beech'). If

"meaning ain't in the head", then no internal check will ensure that I am using terms uniformly. I might try to apply modus ponens and fail.

Circumstances beyond my control can then prevent me from using valid inference forms as touchstones of rationality. This risk of equivocation seems to open the door to skepticism about my own rationality.

Semantic internalists invite us to close the door by affirming the transparency of content (Boghossian 1994). In addition to insulating us against skepticism, transparency explains the asymmetry in rationality between incoherence and ignorance. If meanings are self-intimating, then I have everything needed to avoid logical error. A precondition of normative appraisal is satisfied: 'ought' implies 'can'. In contrast, when I am ignorant about an external fact, no amount of a priori reflection could have helped, so the mistake cannot be blamed on my negligence. I am thereby safe from censure. "Absolutely safe."

## **2. Intrinsic vs. extrinsic equivocation**

All textbook examples of equivocation feature an internal psychological failure that leads to an unwitting shift of meaning. For instance, Irving Copi (1982, 121) opens his discussion of fallacies of ambiguity with a traditional equivocation on 'end': "The end of a thing is its perfection; death is the end of life; hence, death is the perfection of life." As Copi

explains, most words have several meanings. No difficulty arises when we keep the meanings separate. "But when we confuse the different meanings a single word or phrase may have, using it in different senses in the same context, we are using it equivocally." (Copi 1982, 120) This purely mental slip depends solely on the intrinsic properties of introspectible intentions, beliefs, and desires. The speaker never needs to use his sense organs to check whether he is equivocating. Outsiders lack his privileged access and so must detect his equivocations a posteriori through indirect means. Psycholinguists focus on reaction time studies (Kess and Hoppe, 1981). But we informally employ a wide variety of techniques: the method of empathy, statistical generalizations, inference to the best explanation, etc.

Ordinary indexicals demonstrate the possibility of extrinsic equivocation.

Here it is hot.

Here it is humid.

Therefore, here it is hot and humid.

Suppose Mr. Move does not realize that he has moved in the course of the argument. He has equivocated because 'here' picks out different places.

Bad luck!

Internalists about meaning may disagree with externalists about which terms are indexical but not about whether there are indexical terms. For instance, internalists are not persuaded by Hilary Putnam's (1975) twin earth thought experiment suggesting that 'water' is indexical -- along with other natural kind terms. But all internalists agree that some terms are indexical. Where would Descartes be without 'I'? This commitment to indexicals is sufficient to generate extrinsic equivocations. Therefore, no matter how strong his introspective faculty, the reasoner cannot tell whether he is equivocating. It is surprising that reasoners would be defenseless in these cases. But it is a disappointment that all parties must accept. The bad news falls harder on the internalist. For he thought himself safe from undetectable equivocations. But if such equivocations are the basis for logical skepticism, the externalist and internalist are equal victims.

Bertrand Russell thought that an ideal language would preclude equivocation by assigning exactly one meaning to each word. A one to one mapping of words to inner concepts stops intrinsic equivocation. But

it cannot stop extrinsic equivocation. External circumstances endow key terms with other meanings. This vulnerability of the ideal language was obscured by Russell's semantic theory. Russell was a phenomenalist (in his logical atomist days). Phenomenalism makes all meaning depend on intrinsic features of consciousness. If sense data are the basic units of meaning, then the self-intimating nature of sense data will guarantee the transparency of content. Given the extra premises supplied by Russell's brand of ideal language philosophy, extrinsic equivocation is indeed impossible. However, those who profess neutrality on the semantics of the logically perfect language (such as Wittgenstein in the Tractatus) cannot rule out extrinsic equivocation.

### **3. Is extrinsic equivocation fallacious?**

Some equivocations are innocent because they do not undermine validity. Consider `man' in the syllogism "All men are mortal. Socrates is a man. Therefore, Socrates is mortal." If one reads `man' in the unmarked sense in the major premise (as meaning human being) and then in the marked sense in the minor premise (as male adult human being), then the argument is valid. If the equivocation runs in the opposite order, from the specific to the generic, then the argument is invalid.

A validity preserving equivocation can still be criticized on pragmatic grounds. In particular, an argument should not mislead us about its logical form. Uniform usage makes it easier to detect whether the argument instantiates a valid argument form. That is why logicians require that an argument be disambiguated before logical techniques are applied. Consequently, there might be a "pragmatic fallacy" even if there is no semantic fallacy in validity preserving equivocations.

Even so, this would not rescue Mr. Move's inference. His equivocation does undermine the validity of his argument. If Mr. Move were to later discover his equivocation, he would repudiate his earlier inference. Spectators might console Mr. Move by pointing out that he did his best given the available information. They might argue that good intentions are all that is relevant to Mr. Move's rationality. These logical Kantians might have us contrast Mr. Move with someone who intrinsically equivocates but does not extrinsically equivocate.

In particular, suppose Mr. Self-cancellation falsely believes that the Canadian numeral system diverges from the American system after a million. (The British system actually does diverge.) Fortunately, Mr. Self-cancellation is tired and forgets to make this bogus dialect adjustment when he reads from the Canadian paper that Canada minted a billion

coins in 1960. Since he has read in an American newspaper that America minted a billion coins in 1960, he infers that an equal number of coins were minted in America and Canada in 1960. His deduction is luckily sound. For his internal errors cancel out. (That is, the omission caused by his fatigue stops his false belief about the Canadian divergence in number systems from affecting his reasoning.)

Intuitively, Mr. Self-cancellation was illogical even though his argument is valid while Mr. Move was logical even though his argument was invalid. This suggests that only the "subjective" validity of one's deductions are relevant to one's rationality. The principle of charity fortifies this impression. The patient logician resists inferring that his logic students are irrational when they adduce invalid proofs and when they reject valid proofs. He instead attributes false beliefs about how the rules of the logical system work. Error arising from ignorance about rules is not irrationality. Mistakes about logical rules no more indict one's rationality than mistakes about the rules of games such as chess.

Likewise, proficiency in formal logic systems makes one no more logical than proficiency with any other system of rules. Or so it seems if we treat the inward turn as something deeper than a diagnostic strategy.

To avoid crude psychologism one must leave room for logical error. But once inference rules are subjectivized, the contrast between the valid and the apparently valid disappears. The specter of psychologism should draw us to distinctions that explain away the attractions of logical subjectivism. One promising candidate can be adapted from ethics (Dancy 1976):

(a) If one believes that one ought to infer p, then one ought to infer p.

(b) It ought to be that if one believes that one ought to infer p, then one infers p.

Statement (a) is logical subjectivism: believing that an argument is valid makes it valid. (Or more plausibly, believing in accordance with special standards -- of the sort Mr. Move satisfies but not Mr. Self-Cancellation.)

Statement (b) also expresses respect for following one's logical conscience but without endorsing the inference. There is only a defeasible connection between an agent's rationality and the validity of his reasoning.

#### 4. Vague ambiguity refutes transparency

Our syntactic bulwark against psychologism can be strengthened by showing that the transparency thesis fails even if restricted to subjective genres of meaning -- narrow content, one's idiolect, speaker meaning.

The counterexamples issue from the vagueness of `ambiguous'. Mr.

Brush argues:

1. There is a house painter.
2. There is a portrait painter.
3. The house painter is not the same person as the portrait painter.
4. Therefore, there are at least two painters.

Mr. Aesthete asks whether `painter' is ambiguous rather than general (for instance, `elephant' is general between Indian elephants and African elephants). Mr. Brush agrees that if there are two senses of `painter', then he equivocated. But Mr. Brush cannot tell whether there are two senses of `painter' as opposed to two kinds of painters. Neither can Mr. Aesthete. Eventually, they conclude that `painter' is a borderline case of `ambiguous'.

Borderline cases are inquiry resistant. So the vagueness of `ambiguity' suggests that sometimes it is impossible for anyone to know whether the speaker attaches the same meaning to each of two words. Notice that the counterexample is independent of semantic externalism: it does not portray `painter' as subject to linguistic division of labor or portray `painter' as having hidden indexicality. The counterexample works even if we relativize `painter' to Mr. Brush's idiolect.

Dummett might reply that his transparency thesis is restricted to the thesis that definite differences in meaning are always detectable by the speaker. When there is only a borderline difference, there is no fact of the matter. Therefore, the speaker is not ignorant of anything. All genuine meaning differences remain transparent.

However, vague ambiguity can strike where there are definite differences in meaning. Shift focus from vagueness concerning whether a word is ambiguous to vagueness as to which ambiguities a word has. Dictionaries agree that `run' is ambiguous but disagree as to which senses it has. Some of these differences are resolvable, but others are indeterminate. Indeterminacy as to just where an ambiguous word is ambiguous generates definite equivocations that lack definite sites of equivocations. Suppose Mr. Nephew has argued as follows:

1. Since Alan is my father's brother, Alan is my uncle.
2. Since Boris is married to my father's sister, Boris is my uncle.
3. Since Carlos is a friend of my family, Carlos is my uncle.
4. Therefore, I have at least three uncles.

Mr. Nephew's American friend says that this argument equivocates on 'uncle'. There is a sense of 'uncle' that means 'brother of a parent or married to an aunt' (used in premises 1 and 2) and a sense of 'uncle' that means 'man who is a friend of one's family' (used in premise 3). Mr. Nephew's African friend agrees that Mr. Nephew has equivocated but locates it elsewhere. His African friend says that there is a sense of 'uncle' that means 'brother of a parent' (used in premise 1) and a sense that means 'man who is a friend of one's family or who is married to an aunt' (used in premises 2 and 3). As it turns out, Mr. Nephew's idiolect has been equally influenced by his upbringing in America and Africa. So although he agrees that he has equivocated, he cannot tell whether the American or African diagnosis is correct. Mr. Nephew concludes that his use of 'uncle' is borderline between being ambiguous American style and being ambiguous African style. Since his use of 'uncle' is ambiguous

either way, Mr. Nephew knows that he switched meanings in the course of his argument. The existence of the equivocation is definite but the nature of the equivocation is indefinite.

The uncle case is presented as a single argument with two steps, either of which might be the equivocation. One could reorganize some of the reasoning as two arguments:

Since Alan is my father's brother, Alan is my uncle.

Since Boris is married to my father's sister, Boris is my uncle.

Therefore, Alan and Boris are my uncles.

Since Boris is married to my father's sister, Boris is my uncle.

Since Carlos is a friend of my family, Carlos is my uncle.

Therefore, Boris and Carlos are my uncles.

Under the American interpretation, the first argument is valid but the second equivocates. Under the African interpretation, the reverse holds.

Given Mr. Nephew's indeterminacy, he will know that one of the arguments equivocates but not which.

If meaning is transparent, then the speaker should be able to pinpoint any definite difference in his meaning. But he cannot, even in principle. Therefore the transparency thesis is false.

## 5. Univocation

The converse of equivocation is univocation -- treating a univocal expression as if it were ambiguous. Saul Kripke's (1976) univocator is Peter who has the background belief that no politician is a musician. At a concert, Peter hears Paderewski play the piano and comes to believe Paderewski is a musician. At a reception, Peter is introduced to the Polish Prime Minister Paderewski. Not realizing that this politician is the pianist, Peter comes to believe 'Paderewski is not a musician'. Hence, Peter will assent to the contradiction 'Paderewski is a musician and Paderewski is not a musician'. Peter will also judge some valid arguments as invalid, for instance, "Paderewski is a politician. Paderewski is a musician. Therefore, at least one politician is a musician."

Kripke's Pierre puzzle features a more complicated univocation that arises from translation. Pierre is a monolingual Frenchman who hears of a beautiful city, Londres, and so sincerely says 'Londres est jolie'. Pierre moves to England and learns English by immersion. Although he does not

realize it, Pierre has moved to London. However, it is a shabby part of London. Pierre sincerely says 'London is not pretty'.

Kripke portrays the Pierre case as stretching the concept of belief to its breaking point. His commentators play into this air of hysteria by refusing to directly answer Kripke's persistent question 'Does Pierre believe that London is pretty?'. The answer is: Yes, Pierre believes that London is pretty. Yes, Pierre believes that London is not pretty. Yes, Pierre believes that London is pretty and that London is not pretty. (Because Pierre assents to the word salad sentence 'Londres est jolie and London is not pretty'.) And yes, Pierre believes that it is not the case that London is pretty and London is not pretty. (Because he dissents from 'London is pretty and London is not pretty'.) All of this follows because Pierre both believes and disbelieves the contradiction 'London is pretty and London is not pretty'.

Could Pierre's beliefs be more inconsistent? Yes. He would be more inconsistent if he had inconsistencies that could be detected from the intrinsic properties of his belief system. As it stands, Pierre is only extrinsically inconsistent. If we confine our report to the beliefs Pierre has in virtue of his narrow psychology, then there will be no inconsistencies. There can be theoretical purposes for restricting

ourselves to this kind of belief report (say when doing a Bayesian assessment of Pierre's subjective probabilities or when doing psycholinguistic research on Pierre). But our belief reports are not wedded to those purposes.

The feeling that the concept of belief is being stretched to its breaking point is due to a failure to draw the intrinsic/extrinsic distinction. This interplays with a second semantic/pragmatic distinction: Although Pierre believes the contradiction 'London is pretty and London is not pretty', it is misleading to report Pierre's belief that way. For he also dissents to that very sentence, and so also believes the negation of that contradiction. When we report beliefs, there is a conversational implicature that the speaker will assent to the very sentence used to report his belief. But this implicature can be cancelled. For instance, we can say 'Lois Lane believes Clark Kent can fly but she would not put it that way'.

Pierre seems logical because his univocation is extrinsic rather than intrinsic. Pierre's belief that London is pretty and not pretty is an extrinsic belief in a contradiction. So he does not have a priori access to it. This answers Kripke's objection to ascribing inconsistent beliefs to Pierre:

We may suppose that Pierre, in spite of the unfortunate situation in which he now finds himself, is a leading philosopher and logician. He would never let contradictory beliefs pass. And surely anyone, leading logician or no, is in a position to notice and correct contradictory beliefs, if he has them. Precisely for this reason, we regard individuals who contradict themselves as subject to greater censure than those who merely have false beliefs. But it is clear that Pierre as long as he is unaware that the cities he calls 'London' and 'Londres' are one and the same, is in no position to see, by logic alone, that at least one of his beliefs must be false. He lacks information, not logical acumen. He cannot be convicted of logical inconsistency: to do so is incorrect. (Kripke 1976, 257)

Kripke's insistence that we have access to all our inconsistencies conflicts with his causal theory of names (Pettit 1984). The causal theory lets us individuate names by historical processes of which speakers are typically ignorant. Usually, our heuristics about the relationship between names

and their bearers will lead us to the correct conclusion. But given the high probability of some improbable events coming to pass, there will commonly be cases in which the speaker has a justified but false belief that two tokens of the same name have distinct bearers.

Pierre's kind of inconsistency is comparable to conflicting de re beliefs. In W. V. Quine's (1956) saga, Bernard Ortcutt is such that Ralph believes he is a spy and believes that he is not a spy. (Ralph has seen Ortcutt behave suspiciously on the beach but does not realize that it is the same man he has seen behave uprightly in town.) Neither Ralph nor Pierre suffer from inconsistency that is as serious as stereotypical inconsistency in which one has all the resources needed to undo the inconsistency.

## **6. Logical skepticism**

The distinction between intrinsic and extrinsic inconsistency disarms Paul Boghossian's (1992, 1994) criticism that externalism leads to logical skepticism. Boghossian presents a Twin Earth thought experiment featuring an "accidental tourist", Peter. After witnessing Pavarotti floating in Lake Taupo, Peter is whisked away to Twin Earth. (Boghossian is assuming Twin Earth is just another planet in our universe, not another

possible world.) Peter does not notice any difference because Twin Earth has the same appearance as Earth. But there is the famous difference that 'water' on Twin Earth picks out a substance XYZ that is slightly different from Earth's H<sub>2</sub>O -- but only at the level of chemical theory. In the beginning, Peter's use of 'water' picks out H<sub>2</sub>O. But eventually it switches to pick out XYZ as Peter becomes assimilated to the dialect of Twin Earthians. According to Boghossian, when Peter is expressing memories, his use of 'water' reverts to the old reference H<sub>2</sub>O. Hence Peter's use of 'water' is ambiguous. In particular, Peter equivocates when he reasons on Twin Earth 'Pavarotti floated in the water of Lake Taupo, The glass in front of me contains water, therefore, Pavarotti once floated in the same kind of substance that fills this glass'. This would be an equivocation that Peter could not in principle detect.

How do you know that you are not in the same predicament as Peter? Most of your inferences involve vocabulary whose meaning (according to the externalist) "ain't in the head". If you cannot exclude the possibility that you have been switched to Twin Earth, you cannot tell whether you are reasoning validly. For you cannot tell whether you are substituting terms uniformly when applying a rule of inference to the

argument. Boghossian characterizes Peter as a converse case of Kripke's Pierre:

Part of the significance of such a case may be described as follows. The case of Pierre shows that externalism may well lead to a situation in which certain inferences are valid, but don't a priori look to be so. The case of Peter, if it worked, would show something perhaps more surprising: that externalism may well lead to a situation in which certain inferences look valid, but aren't. (Boghossian 1992, 39)

In my terminology, the accidental tourist Peter is just an exotic extrinsic equivocator. Any extrinsic equivocation trivially illustrates the possibility of an invalid argument that looks a priori valid to the equivocator. And any extrinsic univocator trivially illustrates the possibility of a valid argument that looks a priori valid to the univocator. Boghossian is mistaken about there being any extra surprise for the externalist. Any mistake about validity (or consistency) precipitates both kinds of error about validity i.e. mistaking a valid argument as invalid and mistaking an invalid argument as valid.

First take the false belief that an argument ' $P$  therefore  $C$ ' is invalid. This supports the false belief that ' $P \ \& \ \sim C$ ' is consistent. This in turn supports the further error that "' $(P \ \& \ \sim C)$  is inconsistent' therefore  $Q$ '" is (vacuously) valid" on the grounds that the premise is impossible. (Any consistent statement  $Q$  is necessarily consistent, so the meta-statement alleging that  $Q$  is inconsistent is itself inconsistent.)

Second take the false belief that ' $P$  therefore  $C$ ' is valid. This leads to the false belief that ' $P \ \& \ \sim C$ ' is inconsistent. This in turn leads one to count "' $P \ \& \ \sim C$  is consistent" as inconsistent. (Any inconsistency is necessarily inconsistent.) Hence, the victim will view "' $\sim(P \ \& \ \sim C)$  therefore ' $P \ \& \ \sim C$ ' is consistent" as invalid on the grounds that a tautology cannot entail a contradiction. However, the argument is valid under our assumption that ' $P$  therefore  $C$ ' is actually invalid. For then the conclusion of "' $\sim(P \ \& \ \sim C)$  therefore ' $P \ \& \ \sim C$ '" is a tautology. Any argument with a tautologous conclusion is valid. Thus the mistaken belief that ' $P$  therefore  $Q$ ' is valid leads to the mistaken belief that "' $\sim(P \ \& \ \sim C)$  therefore ' $P \ \& \ \sim C$ ' is consistent" is invalid.

Since belief in a contradiction commits one to everything, a mistake about validity commits one to everything. For any mistake about validity can be converted into belief in a contradiction. So although Boghossian's

accidental tourist makes the interesting point more directly and colorfully, Kripke's original case suffices to show that the externalist is committed to innocent acceptance of arguments that are actually invalid.

### **7. A posteriori influences on what is a priori**

Although there is no purely deductive remedy for linguistic misalignments, there commonly are inductive preventions and cures. A mail man can be blamed for sending a package to London, England when it should go to London, Ontario. He is expected to prevent extrinsic equivocation by noting which cities have misleadingly similar names.

Induction also lets us recover from misalignment. As our univocator has more contact with Paderewski, he will have more grounds to doubt his assumption that 'Paderewski' is not a single name. (The politician plays piano at the reception, others say that the politician is a musician, etc.) Duhem's thesis states that any proposition can be consistently held in the face of new evidence by changing background assumptions. Accordingly, our univocator can "consistently" preserve the assumption by making more and more elaborate revisions to his factual beliefs. However, these revisions eventually become dogmatical.

Evolution gives us reason to expect that everyone is open to linguistic correction. Since everybody is always linguistically misaligned about something and alignment has been important as long as human beings have had language, all people ought to be sensitive to empirical evidence that they have mislearned the language. This readiness to change linguistic beliefs is affected by how frequently one discovers errors. Once the linguistic change is made, one counts different sentences as analytic and hence one revises opinions as to which sentences are a priori. For logical purposes, it is convenient to treat reasoners as linguistically perfect. However, this idealization is not suited for individuals who are mistaken about features of their language, such as Pierre. What is a priori need not itself be a priori.

### **8. Blaming policies**

So far, I have discussed extrinsic fallacies in terms congenial to an internalist. I have not questioned the assumption that one is rational as long as one is doing the best one can from one's own perspective. However, an externalist should be inclined to take a more social stand, one that does not confine rationality assessments to the individual's narrow psychology.

Extrinsic equivocations and univocations lower the reliability of one's reasoning. We should expect others to disapprove of these linguistic misalignments because unreliable conclusions tend to be passed along to others. Criticism is punishment. The group achieves a great economy of discipline when it socializes individuals into self-punishment. Just as the immoral come to be pricked by their conscience, the illogical become pricked by logical remorse. These internal sanctions will be colored by social interests. They encourage the individual to play a role in the group's intellectual economy.

Individualists take norms of rationality to be egoistic; only the individual's intellectual self-interest counts. However, groups that had egoistic norms would be outperformed by groups that had norms that sometimes gave preference to the group's cognitive well-being. This epistemic altruism is directly expressed in our intrinsic concern about what others believe and how they came to believe it. Much of the time we spend giving away information and correcting the errors of others could be spent improving our viewpoint. The individual rationally foregoes the opportunity to improve his own rationality because he cares more about an increment in the rationality of other people.

The internal perspective may be so compelling because we confuse conditional rationality judgments with absolute rationality assessments. After the former mathematician, Theodore J. Kaczynski confessed to being the Unabomber, many who conceded that he was rational given his luddite assumptions still denied he was rational. His high degree of logical consistency and rigor were taken as a symptom of mental illness. (One psychiatrist diagnosed Kaczynski as a high functioning paranoid schizophrenic.)

Elsewhere (Sorensen 1991) I have argued that `rationality' is the absence of irrationalities just as `health' is the absence of disease. Some irrationalities do not reduce to bad inferences -- though my guess is that all irrationalities reduce to inefficiencies in getting truth and avoiding error. Opportunity costs that accrue from extreme punctiliousness about one's logic make some valid argumentation irrational. Think of the intricate deductions performed by religious zealots. Logic is a magnet for rule worshippers and ritualists. Nevertheless, the validity of inferences is a central part of rationality. Thus there are extrinsic irrationalities just as their extrinsic diseases such as dwarfism (the relational disease of small stature).

Rationality is unconditional rationality. Research on rationality is facilitated by a piecemeal approach in which we bracket away nuisance factors. For example, Bayesians have made strides by bracketing away the question of how one obtains prior probabilities. This progress tempts them to go beyond mere bracketing and identify rationality with the process of revising probabilities in light of new information. Rationality becomes analogous to pure procedural justice. These Bayesians regard prior probabilities as beyond rational discussion just as many economists view our ends as expressions of taste. For these thinkers, all rationality is about the most efficient means to a goal -- the goal itself drops out.

Likewise, internalists identify rationality with what should be done relative to a perspective. They reject strict liability in favor of 'ought' implies 'can'. Trying to get to the truth takes precedence over getting to the truth. This echoes the pragmatic paradox of Kantian ethics (Sorensen 1995). If the only intrinsically good thing is a good will, trying to do the right thing becomes identical to doing the right thing. But how can the trying itself be what I aim at?

The utilitarian apportions blame in accordance with the consequences of blaming. These consequences include both the general and the special deterrence value of blame plus the operational costs of

meting out blame. For the most part, blame should not be assigned to those who are following their conscience. Still, utilitarianism demands flexibility. Some of its exceptions conform to common sense. We punish moral fanatics who do great harm even when we are persuaded that they are following their conscience.

Our inferences have consequences and so fall under the scope of moral assessment. Hence contradictory beliefs should fall into three cases. First is the familiar class of avoidable contradictions. Second are contradictions that are rendered guiltless by their unavoidability. The final class are unavoidable contradictions that we will nevertheless condemn -- instances of strict liability in logic.

History can vindicate one's logic just as it can vindicate one's morality. At some point, people must just commit to patterns of reasoning and hope for the best. Georg Cantor had difficulty publishing his groundbreaking diagonal argument demonstrating that there are more real numbers than natural numbers. This pattern of reasoning bears a worrisome resemblance to the liar paradox. Editors were concerned that the unorthodox argument was somehow fallacious. Cantor's diagonal argument eventually became an entrenched landmark of mathematical reasoning. But it is not as if Cantor was in a position to predict that his

argument would pass the test of time. Other liar-like aspects of his reasoning in set theory were indeed fallacious -- witness Russell's paradox of the set that contains all and only those sets that do not have themselves as a member.

The cogency of one's reasoning is affected by the social factors outlined by Tyler Burge (1979, 1986) . This connection is not restricted to one's contemporaries because intellectual communities refine concepts over time. Consider the history of arguments employing infinitesimals. The pioneering inferences of Isaac Newton and Gottfried Leibniz looked so illogical that, in 1734, George Berkeley was able to argue that religious reasoning was no less logical than scientific reasoning. The full title of Berkeley's monograph is The Analyst, Or a Discourse Addressed to an Infidel Mathematician. Wherein It is examined whether the Object, Principles, and Inferences of the modern Analysis are more distinctly conceived, or more evidently deduced, than Religious Mysteries and Points of Faith. `First cast out the beam in thine own Eye; and then shalt thou see clearly to cast out the mote out of thy brother's Eye'. Berkeley's allegation of irrationality was no longer contested after Karl Weierstrass developed epsilon-delta style proofs around 1850. These yield the same conclusions (though more clumsily) without relying on

infinitesimals. Mathematicians just conceded that early calculus arguments were riddled with just the sort of incoherencies that Berkeley detailed. Between 1920 and 1950 there were great developments in mathematical logic, in particular, model theory. By 1961, Abraham Robinson was in a position to make sense of infinitesimals with non-standard arithmetic. The pioneers of calculus now look less illogical. But it is anachronistic to say that Robinson was merely articulating something implicit in the thinking of the pioneers. Robinson's proofs are twentieth century proofs. They rely on concepts that were unavailable to eighteenth century thinkers.

Objections to negative numbers, imaginary numbers, and transfinite numbers were rendered fallacious by conceptual developments occurring many generations after the death of the objectors. For instance, Galileo presented a proof that there are as many squares as there rational numbers by putting them into a one to one correspondence with the naturals.

Squares: 1, 3, 9, 16, . . . .

Naturals: 1, 2, 3, 4, . . . .

However, he did not believe his own proof. All squares are naturals but not vice versa. How can a part be as big as the whole? Centuries later,

Richard Dedekind undercut this objection by defining infinite sets in terms of just this property.

Historians of ideas caution against applying contemporary standards and concepts to past thinkers. But this caution often escalates into the historicist position that the future is irrelevant; each thinker must be judged by the standards of his times. But if we insulate past thinkers against judgment of history, we fall into a form of relativism and make nonsense of intellectual progress. Historicism exchanges an insensitivity to the past with an insensitivity to the future.

We measure rationality partly by the extent one is susceptible to criticism about one's inferences. Thus recognition of "strict liability" in logic implies that an individual's rationality is partly a matter of luck. This dovetails with Bernard Williams' thesis that moral status is partly a matter of luck. Indeed, it becomes tempting to unify the two theses. Perhaps one form of luck can be reduced to the other. For instance, logical luck might be a corollary of an ethics of belief that countenanced moral luck. Or, to revert to the rationalist suggestion, moral luck might be reduced to logical luck via a reduction of morality to rationality.

A third possibility is that there is a common cause behind both moral luck and logical luck. Both irrationality and immorality are

connected to certain forms of criticism. The critic usually must respect 'ought' implies 'can'. However, the depth of this principle is over-estimated. The logic of blame sometimes permits censure of the helpless.

Utilitarian blame is sometimes harsher than Kantian blame. But sometimes it is more lenient. The utilitarian attaches no special importance to a priori error. His emphasis on cost-benefit analysis lets him easily accommodate the possibility that there are supererogatory deductions. That is, there are deductive omissions and mistakes that should not be blamed. Euclid's proof that there are infinitely many prime numbers is a priori. But it is pedantic to blame people for failing to draw the inference. The fact that they have all the resources needed to make the discovery is no more telling than the fact that a pharmacist has all the resources in his shop to cure currently incurable diseases. Problem solving requires skill as well as material. Better consequences flow from undertaking cheap empirical investigation rather than costly reflection. This explains the irrationality of adding sums mentally instead of using a calculator.

Thus the utilitarian is poised to admit that some avoidable contradictions are also blameless. Exactly which will be sensitive to skills and social circumstances. A mathematician is more embarrassed by

mathematical mistakes than a layman. For there is a point to blaming those who could have easily avoided the error and rarely any point in blaming the intractably incompetent. Thus logical luck displays contextual subtleties reminiscent of Aristotle's doctrine of the mean.

\* Improvements of this paper can be traced to listeners at Uppsala University, University of Adelaide, University of Queensland, University of Canterbury , and at a session of the Australasian Association Philosophy (New Zealand Division) at Massey University. I am especially grateful for Andre Gallois' suggestions.

## REFERENCES

- Boghossian, Paul (1992) "Externalism and Inference" Rationality in Epistemology ed. Enrique Villanueva (Atascadero, California: Ridgeview Publishing) 11-28.
- \_\_\_\_\_ (1994) "The Transparency of Content" Philosophical Perspectives 8 ed. James Tomberlin (Atascadero, California: Ridgeview Publishing Company).
- Burge, Tyler (1979) "Individualism and the Mental" Midwest Studies in Philosophy 4: 73-121.
- \_\_\_\_\_ (1986) "Intellectual Norms and Foundations of Mind" The Journal of Philosophy vol. LXXXIII no. 12, 697-719.
- \_\_\_\_\_ (1988) "Individualism and Self Knowledge" Journal of Philosophy 85: 649-63.
- Butler, Keith (1997) "Externalism, Internalism, and Knowledge of Content" Philosophy and Phenomenological Research LVII/4: 773-800.
- Copi, Irving (1982) Introduction to Logic (New York: Macmillan Publishing Co) Sixth Edition.
- Dancy, Jonathan (1976) "The Logical Conscience" Analysis 37/2: 81-84.
- Dummett, Michael (1978) "Frege's Distinction between Sense and Reference" Truth and other Enigmas (Harvard: Harvard University Press).

Falvey, K. and Owens, H. (1994) "Externalism, Self-Knowledge, and Skepticism" The Philosophical Review 103: 107-37.

Kess, Joseph F. and Hoppe, Ronald A. (1981) Ambiguity in Psycholinguistics (Amsterdam: John Benjamins).

Kripke, Saul (1976) "A Puzzle about Belief" in Meaning and Use ed. Avishi Margalit (Dordrecht: D. Reidel)

Ludlow, Peter (1995) "Externalism, Self-Knowledge, and the Prevalence of Slow Switching" Analysis 55/1: 45-49.

McKay, D.. G. and Bever, T. G. (1967) "In search of ambiguity" Perception and Psychophysics 2: 193-200.

McKinsey, Michael (1991) "Anti-Individualism and Privileged Access" Analysis 51: 9-16.

Monk, Ray (1990) Ludwig Wittgenstein: The Duty of Genius (New York: The Free Press).

Nagel, Thomas (1979) "Moral Luck" Mortal Questions (Cambridge University Press).

Pettit, Philip (1984) "Dissolving Kripke's Puzzle about Belief" Ratio XXVI 2: 181-193.

Putnam, Hilary (1975) "The Meaning of `Meaning'" Language, Mind and Knowledge (Minneapolis, Minnesota: University of Minnesota Press) 131-193.

Quine, W. V. (1956) "Quantifiers and Propositional Attitudes" The Journal of Philosophy 53: 177-187.

Sorensen, Roy (1991) "Rationality as an Absolute Concept" Philosophy 66/258: 473-486.

\_\_\_\_\_ (1995) "Unknowable Obligations" Utilitas 7/2: 247-271.

Warfield, T. (1992) "Privileged Self-Knowledge and Externalism are Compatible" Analysis 52: 232-237.

Williams, Bernard (1976) "Moral Luck" Proceedings of the Aristotelian Society supplementary volume 1: 115-135.