

## REWARDING REGRET

Announcement: You are hereby offered a choice between \$1 and \$10. In addition, there is a bonus of \$100 if you regret your choice.

### I. MAXIMIZING REGRET

You are tempted to choose \$1 in the hope of regretting this choice. But this is futile. You cannot regret a decision that you perceive as maximizing your overall pay-off.

You contemplate a crazed grab for \$1 in the hope that this will stimulate self-reproach. Your mother is supportive. She is a world authority on guilt. She describes psychological studies (such as those listed for chapters 2 and 7 of Landman 1993) demonstrating that regret is sharpest for unusual choices that could have easily been decided the other way. However, your lawyer notes that the regret must concern what you chose rather than how you chose it. And the regret must last through a cool hour of reflection. If, at the end of the hour, you regret that you chose \$1, then our eccentric hundredaire immediately gives you \$100 cash and walks out of your life. Any regrets or changes of mind after the hour are irrelevant. Your mother consoles you by noting that regret also correlates with the imminence and irreversibility of the outcome.

Your anesthesiologist suggests that he induce amnesia after you choose \$1. You would remember choosing \$1 rather than \$10. But you would have forgotten that you did this for the sake of the \$100 bonus. Thus you would have regret without irrationality. Alas, your lawyer draws attention to the fine print of the offer which rules out disruptions of your knowledge. This also prohibits you from arranging to be deceived or having your inferential abilities distorted. (As in the surprise examination paradox, there is no problem in the announcement being true, only in you knowing that the announcement is true.)

Your family hypnotist suggests a more lucrative plan that will not violate the epistemic codicils. Choose \$10. The hypnotist will then implant regret. However, your lawyer cautions that the offer requires that you believe that you made the wrong choice. Mere feelings of regret do not pay-off.

On the other hand, there is no requirement that your regret be well founded. The offerer does not care whether you actually make the wrong choice. You need only wish that you had chosen differently. That wish could be based on a variety of beliefs. Perhaps you believe your choice was irrational (Gaifman 1983). Maybe you merely believe that the other option would have been more fruitful. Whatever. Only the regret counts, not its pedigree. A regretter receives the bonus even if his ancillary beliefs are false, unjustified, or defective in yet subtler ways.

## II. A CONTRAST WITH UNSTABLE CHOICE

At this juncture one may just conclude that the \$10 choice is best. After all, the \$10 choice dominates the \$1 choice:

	Regret	No Regret
Choose \$1	\$101	\$1
Choose 2	\$110	\$10

Appeals to dominance are questionable when the choice influences the "state of nature". For instance, I ought not to argue for smoking on the grounds that it dominates abstention:

	Cancer	No Cancer
Abstain	Fourth Best	Second Best
Smoke	Third Best	Best

Smoking increases the risk of cancer. Yet there may be a redeeming difference when the dominance principle is applied to the regret puzzle.

There is no way to affect your subjective probability of regret. Hence, one might argue that the independence restriction is satisfied.

Nevertheless, this argument (or indeed any argument) in favor of choosing one alternative is self-defeating. If you show that one alternative is better, then you thereby show that you should choose the opposite. This makes the regret puzzle resemble an "instability situation" in which one will regret whichever choice one makes. For example, Brian Skyrms (1982) supposes that a mean demon has put a prize in one of two boxes, specifically, the box you will not choose. You will regret either choice.

But if regret were inevitable in the regret puzzle, you would rejoice; the bonus would be inevitably won. And given that the bonus is won either way, you should take the \$10 rather than \$1. You would advise your friend who has been given the same offer to take \$10 and criticize him if he planned to take \$1. But your warning is contradictory: "If you choose \$1, you will regret it but if you choose \$10 you will be glad you did and regret that you did".

Perhaps the warning can be rendered consistent by defending the co-possibility of regret and rejoicing. You invest in your wife's education to ensure that she gets a high paying job. She winds up with a higher salary than yours. You rejoice in the extra income but regret your loss of status as primary wage earner. Your ambivalence about the investment is consistent if you are merely lamenting some aspects of a choice while

being glad about other aspects. For this restricted regret is compatible with endorsing the investment decision overall.

In the regret puzzle, one must repudiate the choice. This is stronger than prohibiting the agent from ratifying his choice (Jeffrey 1981). The bonus can only be won by doing something common sense and many philosophers (Nagel 1970, Rawls 1971) condemn as irrational: choosing A over B when you know you will regret choosing A over B. Indeed, anyone who plans to secure the bonus must do something even worse: he must intend the regret, not merely foresee it. For the regret plays an essential role in securing the bonus; it is not an epiphenomenal residue of a past choice or an act of self-punishment aimed at improving future choices. Instead of being a side-effect (as in instability scenarios), the regret is functional -- an indispensable means to an end.

### III. THE SCOPE OF REASONS

A friend of the dying W. C. Fields was surprised to find him reading the Bible in his hospital bed. Fields, a notorious libertine, explained: "Just looking for loopholes."

The New Testament emphasis on an all-forgiving God might be just what W. C. Fields was looking for. For it suggests the following scheme: enjoy sinful pleasures, repent, then be admitted to Heaven. Some Christians have tried to close this loophole. Some point to prudential

drawbacks such as the risk of a sudden death. But many ascribe a logical flaw: the plan is incoherent because regret is not voluntary. I can say 'I regret it' on a whim but I cannot regret on a whim. My reason for saying 'I regret it' can be a reward for uttering those words. But my reason for regret cannot be that I will be rewarded for the regret.

A (misleadingly) similar objection was lodged against Pascal's Wager. Showing that theism is a wiser investment than atheism only gives the atheist a practical reason to believe that God exists. Belief is only sensitive to epistemic reasons.

The reasons in the regret puzzle are all practical. Examples of epistemic reasons for

(W) Governor George Wallace regrets blocking black students from entering the University of Alabama.

dwell on signs of regret: Wallace's apologies, his increasingly vigorous acts of atonement, his pained preoccupation with the past. Psychological research on regret (summarized in Landman 1993) adds plausibility to the hypothesis that Wallace did indeed regret. Regret correlates with suffering, intimations of mortality, and leisure for reflection -- all caused by an assassination attempt that left Wallace paralyzed from the waist down. The narrow scope, practical reason for (W) is Wallace's judgment

that he was wrong to block the students. Speculation on the wide scope practical reasons for Wallace's regret focus on benefits that accrue from regret: forgiveness from blacks, relief from pangs of conscience, rejuvenated political prospects. Interestingly, there are no complete, purely epistemic reasons for the object of regret. Showing that p is true only satisfies a necessary condition for regretting p. The best one can do is to supply reasons that conjoin facts and values. For instance, some opponents of Hume's fact/value gap portray desire as evaluative belief. They might say Wallace regrets a "moral fact", namely that Wallace wrongly blocked black students from entering the University of Alabama.

The distinction prompted by the regret puzzle is syntactic rather than semantic:

Narrow scope: ( q)(I regret that p & REASON(q, p))

Wide scope: ( q)(I regret that p & REASON(q, I regret that p))

Regret that p is only sensitive to narrow scope reasons i.e. reasons bearing on p, rather than reasons bearing on regret that p. The regret puzzle shows that a propositional attitude may be sensitive to the scope of reasons, not just the sense of 'reason' or species of reason.

The scope distinction can be drawn within the realm of epistemic reasons. 'There is a reason for the Pope believing that God exists' has a

narrow scope epistemic reading (where, say, the Pope bases his belief on the premises of the Design argument) and a wide scope epistemic reading (where, say, the Pope's theism is evident from the surveillance equipment that has been installed in the Vatican). And of course, there is the third reading in which 'reason' is read prudentially: 'The Pope would lose his job if he did not believe that God exists'. The point is that I am not claiming to discover a new sense of 'reason'. I am drawing attention to syntactic ambiguities that have motivational asymmetries that parallel those found for the different senses of 'reason'.

The syntactic formulation exposes more structure than the informal technique of emphasizing key words. Logical grouping indicators help us import lessons learned from modal logic and smoothly generalize the distinction between narrow and wide scope reasons. For instance, the reason might support a range of propositional attitudes rather than a particular attitude. Buridan's ass has a reason to make some choice or other but no reason to choose the left bale and no reason to choose the right bale:

$(\ q)REASON(q, \text{Choose left or Choose right}) \ \& \ \neg(\ q)(REASON(q, \text{Choose left}) \ \& \ \neg(\ q)REASON(q, \text{Choose right}))$

This brings out a neglected resemblance between the ass and French existentialists. Jean Paul Sartre gives reasons to take a stand on certain issues but at the same time gives no reason to take one stand rather than another. There are reasons to be committed but no reason to be committed to this or that.

Oh very well, the American pragmatists also resemble the ass. In "The Will to Believe", William James promotes selective leaps beyond the evidence. But he is skeptical about there being reasons for leaping in one direction rather than another. The primary thrust of James' essay is that (wide scope) practical reasons are sometimes a legitimate basis for adopting a belief. However, there is also a secondary theme of discretion. James forces us to choose but hopes to preserve freedom of choice by denying that there is a reason for one particular option over its rivals.

#### IV. ITERATED REGRET

Spinoza maintained that one never has reason to regret: "Repentance is not a virtue. i.e. it does not arise from reason. Rather, he who repents what he did is twice miserable, i.e. impotent." (Ethics pt. IV, prop. 54)

Spinoza must have struggled with regrets that arose through force of habit. This struggle must have been marked by faltering steps in the direction of an infinite regress: Spinoza regrets and then regrets having

broken his policy against regrets -- creating yet another regrettable violation.

A more restricted policy against regret avoids the regress. An admiral can stably regret his infraction of his personal ban against regrets concerning naval decisions. The admiral may justify the ban on the grounds that regrets erode the martial virtue of decisiveness.

Philosophers have more generic reasons for rejecting regret: the pain of regret is unnecessary and fosters a useless preoccupation with the past. Instead of letting others socialize your conscience into an instrument of self-punishment, you should turn your back on sorrow and boldly stride into the future.

Notice that these are wide scope reasons against regret. Only nihilists attack regret by citing narrow scope reasons; they argue that since nobody does wrong, you did not do wrong. In contrast, moral realists (Bittner 1992) oppose regret in the way a utilitarian might oppose punishment. Since punishment is intrinsically bad, the utilitarian can only justify punishment as a necessary evil. If the utilitarian doubts that punishment (sufficiently) deters, incapacitates, or reforms, then he will oppose punishment as a useless heaping of evil upon evil. This utilitarian would still believe that people act wrongly. But he would deny that wrong-doers should be punished. He might even deny that they should be blamed. He is then close to joining the opponents of regret.

Historically, utilitarians have been optimistic about the consequences of punishment and would fashion those reasons into a rationale for regret. Biologists would also be quick to suggest reasons for regret: we evolved regret for much the same reason we evolved pain. Just as pain directs attention to its possible causes, regret brings to mind counterfactual outcomes. These unpleasant thought experiments prepare the agent for better decisions. Regret also has reputation effects that lead these self-punishers to outperform those who directly maximize their interests (Frank 1988). Once again, these are wide scope reasons for regret. Wide scope reasons have no power to make one regret or refrain from regret. They can only support indirect means of controlling regret.

When regret embeds another regret (or absence of regret), we must relativize the narrow/wide distinction to the two attitudes. A mercy killer realizes that he will be sentenced leniently if he regrets the killing and harshly if he does not. The mercy killer morally prefers leniency but believes the killing was justified. So although he does not regret the killing he does regret that he does not regret the killing.

( q)[-REASON(q, - He killed) & REASON(q, (Regretting that he killed))]

Notice that one quantifier governs the whole formula. A single reason, leniency, does double-duty. Leniency is only a wide scope reason for

regretting the killing and so cannot ground that regret. But leniency is a narrow scope reason for regretting that one does not regret killing.

Therefore, leniency does ground the meta-regret.

Iterated regret is inconsistent when the meta-regret has the same level of reason as the object level regret. In other words, regret that p clashes with meta-regret that p when both are based on reasons concerning p itself. For the regret must accept the reason while the meta-regret rejects the reason.

To see this, first note that 'regret' is truth entailing: regret that p entails p. 'Nietzsche regrets that he never regrets' is as contradictory as 'Socrates knows that he knows nothing.' One cannot paraphrase 'regret that p' as 'wishing that p did not happen' because wishing is not truth entailing.

'George Bush regrets that he raised taxes' entails 'George Bush raised taxes'. Therefore, higher order regret entails lower order regret: 'George Bush regrets that he regrets that he raised taxes' entails 'George Bush regrets that he raised taxes'. But if the regret and meta-regret are both motivated by reasons governing 'George Bush raised taxes', then Bush repudiates raising taxes and repudiates that repudiation. Repudiating one's current repudiation is as absurd as saying 'I am sorry but I take back this very apology'. I can only retract a pre-existing apology. An aborted apology is not an apology. Iterated regret is only

coherent when scope distinctions prevent the meta-regret from endorsing what the regret repudiates.

#### V. COMPARISON WITH INTERNAL/EXTERNAL REASONS

Talk of grounding brings to mind Bernard Williams' (1985) distinction between internal and external reasons. A pacifist in a family with a military tradition has an external reason to choose a military career but not a reason he internalizes. The reason cannot be used to explain or justify his choice. Similarly, there is an external reason for an amoralist to be altruistic but not a reason that moves him. Although all of Williams' examples concern narrow scope reasons, his internal/external distinction can be applied to wide scope reasons. If the amoralist is informed that a million dollars will go to the poor if he prefers positive eugenics over negative eugenics, then he has an external, wide scope reason to prefer positive eugenics. The preference is doubly unmotivated. The same offer to a moral person would be an internal, wide-scope reason. The preference is ungrounded again but only by virtue of scope.

The internal/external distinction is drawn within the realm of practical reasons. External epistemic reasons are impossible because an agent who recognizes a reason as epistemic thereby manifests a readiness to change his belief in accordance with the reason. External practical reasons are possible because I can recognize that a reason can make an

option desirable without thereby manifesting any personal interest in that reason.

Williams' point is that only a proper subset of practical reasons motivate. Instead of distinguishing between two senses of 'reason', he is distinguishing between two species of practical reason: internal and external.

The internal/external distinction is based on an individual's particular motivational psychology rather than a syntactic feature of reason sentences. The scope distinction mimics the motivational asymmetry of the internal/external distinction because propositional attitudes have intrinsic aims and aversions -- "formal objects" in medieval terminology. Attitudes that aim at truth are only sensitive to narrow scope reasons. Noncommittal attitudes (suppose, entertain, imagine) are sensitive to wide scope reasons. If offered \$100 to think about the price of eggs in China, you can earn the money.

## VI. TOXIN AND THE SYNTAX OF SELF-EFFACEMENT

Some attitudes commit to the epistemic possibility of their propositional objects rather than their truth. If I know that something won't happen, I cannot intend that it happen. This principle lies behind Gregory Kavka's toxin puzzle. An eccentric billionaire has promised you a million dollars if, at midnight, you intend to drink a vial of toxin the next day. Drinking

the toxin would make you ill for a day. That is bad but you would gladly suffer that amount of misery to become a millionaire. However, the billionaire emphasizes that the million dollars will be awarded for intending to drink the toxin. Whether you actually drink the toxin is irrelevant. Kavka concludes that you cannot make the million because you know that you will not drink the toxin.

The toxin puzzle grew out of Kavka's (1978) interest in the paradoxes of deterrence. In a policy of mutual assured destruction, each party has a reason to intend to retaliate if attacked but no reason to retaliate if attacked. Intending to retaliate has an instrumental justification as deterrence. But once the deterrence fails, it is pointless to follow through on the threat.

David Gauthier (1994) claims that his theory of rationality solves the toxin puzzle along with others such as the prisoner's dilemma and Newcomb's problem. His basic premise is that the rationality of a plan is divisional: all of the components of a rational plan are themselves rational. This allows him to commend follow-ups that do no good in themselves. After collecting the million, Professor Gauthier drinks the toxin. After President Gauthier fails to deter the enemy through mutual assured destruction, he initiates that massive second-strike he has been threatening.

Orthodox decision theorists can only justify actions by their effects. This causal requirement makes time relevant. Since causes precede effects, an action cannot be justified by what preceded it; only by what lies ahead. Gauthier's constrained maximization is Janus faced because it lets mereological relations justify. If an action is part of a rational plan, then rationality flows top-down to the action itself. The temporal position of the part is irrelevant. In contrast, a straightforward maximizer cannot bring himself to drink the toxin after he has collected his million. He cannot push the button after the enemy has delivered a devastating first strike. Each situation is judged afresh. The straightforward maximizer refuses to be bound by the "dead hand of the past".

Gauthier's holistic theory solves the toxin puzzle by permitting retrospective justification: the reason Gauthier is now drinking the toxin is that it is part of a rational plan he inaugurated yesterday. The rationality of the plan derives from the classic maxim of maximizing expected utility. Gauthier's innovation is the requirement that the maxim only apply to the big picture.

However, holism does not help one earn the bonus for regret. Gauthier cannot condemn his past choice on the grounds that this disownment is part of his overall scheme for maximizing expected utility. For if Gauthier regards the choice as part of the rational plan, he must regard it as rational and thereby not regret it.

The regret puzzle undermines a similar solution to the toxin puzzle propounded by Edward McClennen (1990, 231). McClennen notes that a resolute agent who can directly commit to a course of action fares better than one who can only indirectly commit by using external gimmicks to bind himself (such as paying a hit man to kill you if you do not drink the toxin). A rational agent will therefore cultivate resoluteness even though this virtue involves sacrifices -- such as drinking the toxin just because you committed to drinking the toxin. McClennen's solution fails to generalize to the regret puzzle because doing what you will regret does not make you more resolute.

So neither McClennen nor Gauthier can claim that their theories outperform the orthodox theory for the regret puzzle. The people who do best are regretters who choose \$10. These are muddled individuals. I don't envy them. But orthodox decision theory says that, for the right price, I should take a pill that would muddle me. Revisionist theory had better agree. Hence revisionists must agree that action guidance is not a criterion of adequacy.

The distinction between wide and narrow scope suggests a way of separating acceptable self-effacement from the vicious variety. It is permissible to give a reason for subscribing to a rival theory but it is impermissible to give a reason for the rival theory. If theory T1 says that theory T2 is more accurate, then the truth of T1 entails its falsehood but

not vice versa. Hence, a subscriber to T1 should switch to T2. However, when T1 says that subscribing to T2 yields better consequences than subscribing to T1, then T1 could still be the accurate theory. Wide scope reasons for self-effacement are permissible, narrow scope reasons are forbidden. Both the orthodox theory and revisionism can be self-effacing given the wide scope reading. There is enough humble pie for everyone.

\*This paper was presented at the Philosophical Society in Lund, Sweden.

Wlodek Rabionwicz re-oriented my thinking about the regret puzzle. I also thank Kent Bach, Eric Barnes, Haim Gaifman, Brian Loar, Gary Malinas, and Graham Oddie.

#### REFERENCES

Bittner, Rudiger (1992) "Is it reasonable to regret?" Journal of Philosophy 262-273.

Frank, Robert (1988) Passions within Reason: The Strategic Role of the Emotions (New York: Norton).

Gaifman, Haim (1983) "Paradoxes of infinity and self-applications, I" Erkenntnis 20: 131-155.

Gauthier, David (1994) "Assure and Threaten" Ethics 104 (July 1994): 690-721.

Jeffrey, Richard (1981) "The Logic of Decision Defended" Synthese 48, 473-492.

Kavka, Gregory (1978) "Some Paradoxes of Deterrence" Journal of Philosophy.

\_\_\_\_\_ (1983) "The Toxin Puzzle", Analysis 43: 33-6.

Landman, Janet (1993) Regret (New York: Oxford University Press).

McClennen, Edward (1990) Rationality and Dynamic Choice (Cambridge University Press).

Mele, Alfred (1996) "Rational Intentions and the Toxin Puzzle" Proto Sociology 8/9: 39-52.

Nagel, Thomas (1970) The Possibility of Altruism (Oxford University Press).

Rabinowicz, Wlodek (1989) "Stable and Retrievable Options" Philosophy of Science 56: 624-41.

Rawls, John (1971) A Theory of Justice (Harvard University Press).

Skyrms, Brian (1982) "Causal Decision Theory" Journal of Philosophy 79/11: 695-711.

Sobel, Howard (1994) Taking Chances (New York: Cambridge University Press).

Weirich, Paul (1981) "A Bias of Rationality" Australasian Journal of Philosophy 59/1: 31-37.

Williams, Bernard (1985) "Internal and External Reasons" in his Moral Luck (Cambridge University Press).