[Published in *Erkenntnis* 75 (1): 147–150, 2011, DOI: 10.1007/s10670-010-9267-6.
The final publication is available at www.springerlink.com.]

A Note in Defence of Ratificationism

Johan E. Gustafsson*

Andy Egan argues that neither evidential nor causal decision theory gives the intuitively right recommendation in the cases The Smoking Lesion, The Psychopath Button, *and* The Three-Option Smoking Lesion. *Furthermore, Egan argues that we cannot avoid these problems by any kind of ratificationism. This paper develops a new version of ratificationism that gives the right recommendations. Thus, the new proposal has an advantage over evidential and casual decision theory and standard ratificationist evidential decision theory.*

Andy Egan argues that neither evidential decision theory (EDT) nor causal decision theory (CDT) gives the right recommendation in the cases *The Smoking Lesion* and *The Psychopath Button*.

> *The Smoking Lesion*
> Susan is debating whether or not to smoke. She believes that smoking is strongly correlated with lung cancer, but only because there is a common cause—a condition that tends to cause both smoking and cancer. Once we fix the presence or absence of this condition, there is no additional correlation between smoking and cancer. Susan prefers smoking without cancer to not smoking without cancer, and she prefers smoking with cancer to not smoking with cancer.[1]

> *The Psychopath Button*
> Paul is debating whether to press the "kill all psychopaths" button. It would, he thinks, be much better to live in a world with no psychopaths. Unfortunately, Paul is quite confident that only a **[p. 148]** psychopath would press such a button. Paul very strongly prefers living in a world *with* psychopaths to dying.[2]

Egan takes the only rational actions in these cases to be smoking in *The Smoking Lesion* and not pressing the button in *The Psychopath Button*. But EDT recommends not smoking and CDT recommends pushing the button. The standard ratificationist version of EDT will not do either, since not pressing the button is unratifiable.[3]

In search of a better theory he considers what he calls 'Lexical Ratificationism':

* I would be grateful for any thoughts or comments on this paper, which can be sent to me at *johan.eric.gustafsson@gmail.com*.

[1] Egan (2007, p. 94).

[2] Egan (2007, p. 97).

[3] See Jeffrey (1983, p. 16).

> An act A is *ratifiable* if and only if there is no alternative B such that $\mathrm{VAL}(B)$ exceeds $\mathrm{VAL}(A)$ on the supposition that A is decided upon.[4]
>
> *Lexical Ratificationism (LR)*: It is rational to perform an act A if and only if:
>
> (1) A is ratifiable and there is no other ratifiable option with higher $\mathrm{VAL}_{\mathrm{EDT}}$ than A, *or*
>
> (2) There are no ratifiable options, and no other (unratifiable) option has higher $\mathrm{VAL}_{\mathrm{EDT}}$ than A.[5]

Although LR gives what Egan takes to be the right recommendations in *The Smoking Lesion* and *The Psychopath Button*, he abandons this proposal since it rules out smoking cigars and smoking cigarettes as irrational on the wrong grounds in Anil Gupta's *The Three-Option Smoking Lesion*.

> *The Three-Option Smoking Lesion*
> Samantha has three options: Smoke cigars, smoke cigarettes, or refrain from smoking altogether. Call these options CIGAR, CIGARETTE, and NO SMOKE. Due to the ways that various lesions tend to be distributed, it turns out that cigar smokers tend to be worse off than they would be if they were smoking cigarettes, but better off than they would be if they refrained from smoking altogether. Similarly, cigarette smokers tend to be worse off than they would be smoking cigars, but better off than they would be refraining from smoking altogether. Finally, nonsmokers tend to be best off refraining from smoking.[6]

Here LR just recommends NO SMOKE because it is the only ratifiable action. However this seems wrong, since if you find yourself deciding on CIGAR or CIGARETTE then you have good reason to think that NO SMOKE is not the way to go. Egan rejects LR because it rules out CIGAR and CIGARETTE in favour of NO SMOKE on grounds of its ratifiability. But he does not just think that Gupta's case is a counterexample to LR: **[p. 149]**

> More importantly, though, this kind of case is fatal for ratificationism in general. *No* ratificationist account will be able to deliver the right results in the sorts of three-option cases that Gupta has pointed out. The real importance of the Gupta cases is not that they refute *lexical* ratificationism—it's that they refute *every* form of ratificationism.[7]

However, Egan does not support this stronger claim with any further argument. I take *The Psychopath Button* and *The Three-Option Smoking Lesion* to just show that standard ratifiability is too strong, since unratifiable options seem rational in these cases. In the following I will propose a weakened version of ratificationism that gets all of Egan's cases right.[8]

[4] Egan (2007, p. 107).
[5] Egan (2007, p. 111).
[6] Egan (2007, p. 112).
[7] Egan (2007, pp. 112–113).
[8] See Weirich (1986, pp. 444–445) and Rabinowicz (1989, pp. 633–634) for two previous weakenings of ratifiability. However, as noted by Rabinowicz, both of their proposals violate strong domination.

> An act A is *generally ratifiable* if and only if there is no alternative B such that for every alternative C, $\mathrm{VAL}(B)$ exceeds $\mathrm{VAL}(A)$ on the supposition that C is decided upon.
>
> *General Ratificationism (GR)*: It is rational to decide upon an act A if and only if A is generally ratifiable and there is no other generally ratifiable option with higher $\mathrm{VAL}_{\mathrm{EDT}}$ than A.

Although GR grants the choice of some unratifiable acts as rational, it rules out as irrational the choice of any act that is dominated by another act in the sense that whatever the agent decides to do he would always be better off with the other alternative. Note also that given that the set of available alternatives is finite, there will always be at least one generally ratifiable alternative; because otherwise, for every available alternative A there would have to exist some further available alternative B such that whatever the agent decides to do B has higher $\mathrm{VAL}_{\mathrm{EDT}}$ than A—thus landing us in an infinite regress. Hence, there is no need for a lexical version of GR.

Like LR, GR only recommends smoking as rational in *The Smoking Lesion* since it is the only generally ratifiable option and it also only recommends not pressing the button as rational in *The Psychopath Button* since both options are generally ratifiable and not pressing has a higher $\mathrm{VAL}_{\mathrm{EDT}}$. But unlike LR, GR does not rule out CIGAR or CIGARETTE in *The Three-Option Smoking Lesion* on grounds of NO SMOKE's ratifiability, since all options are generally ratifiable. If CIGAR or CIGARETTE is ruled out in favour of NO SMOKE it is due to a higher $\mathrm{VAL}_{\mathrm{EDT}}$ of NO SMOKE. Thus, GR handles adequately all the cases above. However, GR is still in need of revision since it yields divergent recommendations in two scenarios presented to me by Frank Arntzenius.

> *Scenario 1*
> You are confronted with three boxes A, B, and C. However, you can only choose A or B in this first scenario. A perfect predictor has filled the boxes with money. If the predictor predicted that you will take A then he filled the boxes as follows: \$2 in A, \$1 in B, and \$0 in C. If he predicted you will take B then he filled the boxes as follows: \$4 in A, \$3 in B, and \$0 in C. **[p. 150]**
>
> *Scenario 2*
> In this second scenario you are also confronted with A, B, and C but this time you may choose any one of them. If the predictor predicted that you will take A or B he filled the boxes as in scenario 1. If he predicted you will take C then he filled the boxes as follows: \$1 in A, \$2 in B, and \$0 in C.

The problem is that GR recommends A in *Scenario 1* but B in *Scenario 2*. The addition of the clearly dominated alternative C in *Scenario 2* should not alter the ranking between A and B. While GR appropriately rules out C since C is not generally ratifiable, it wrongly ignores that B is not generally ratifiable in a choice between the remaining options A and B. This problem can be overcome if we adopt an iterated variant of GR.

> An act A is *generally ratifiable*$_0$ if and only if there is no alternative B such that for every alternative C, $\mathrm{VAL}(B)$ exceeds $\mathrm{VAL}(A)$ on the supposition that C is decided upon.
>
> An act A is *generally ratifiable*$_{n+1}$ if and only if there is no generally ratifiable$_n$ alternative B such that for every generally ratifiable$_n$ alternative C, $\mathrm{VAL}(B)$ exceeds $\mathrm{VAL}(A)$ on the supposition that C is decided upon.
>
> An act A is *generally ratifiable*$_*$ if and only if for all $k \geq 0$, A is generally ratifiable$_k$.
>
> *Iterated General Ratificationism (IGR)*: It is rational to decide upon an act A if and only if A is generally ratifiable$_*$ and there is no other generally ratifiable$_*$ option with higher $\mathrm{VAL}_{\mathrm{EDT}}$ than A.

IGR gives the same recommendations as GR in all the cases above except for *Scenario 2* where it recommends A. GR's recommendation, B, is ruled out by IGR since it is not generally ratifiable$_*$. Although both A and B are generally ratifiable$_0$, only A is generally ratifiable$_1$. Thus, IGR gives the same recommendation in both scenarios. Furthermore, like GR, IGR is a theory that gives the intuitively right recommendations in Newcomb cases like *The Smoking Lesion*, cases like Egan's *The Psychopath Button*, and Gupta cases. This gives IGR an advantage over Egan's LR and the main contenders EDT, standard ratificationist EDT, and CDT.

Thanks to Frank Arntzenius, John Cantwell, Sven Ove Hansson, Martin Peterson, Wlodek Rabinowicz, Tor Sandqvist, and two anonymous referees for *Erkenntnis*.

References

Egan, Andy (2007) 'Some Counterexamples to Causal Decision Theory', *The Philosophical Review* 116 (1): 93–114.

Jeffrey, Richard C. (1983) *The Logic of Decision*, Chicago: University of Chicago Press, second edn.

Rabinowicz, Wlodek (1989) 'Stable and Retrievable Options', *Philosophy of Science* 56 (4): 624–641.

Weirich, Paul (1986) 'Decisions in Dynamic Settings', in *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pp. 438–449, Philosophy of Science Association.