

Sider's stage theory and expectancy prior to personal fission

According to Sider's stage theory a subject about to undergo personal fission should expect to experience each outcome simultaneously as distinct persons. How is the subject to make sense of this? I argue that their most paradigmatically self-interested future-directed behaviour, betting for personal gain, ought to be exactly the same as in equivalent games of chance where the possible outcomes correspond to the fission output branches. So this novel form of expectancy, albeit strange, can be a reliable guide to action.

The concept of personal fission has had a significant place in discussions of personal identity, especially in the renowned work of Derek Parfit (1984), but the idea is very bizarre so the motivation can seem thin for concern about what a person ought to expect if faced with the prospect of fission. Still, fission may well be physically possible and it certainly seems to be imaginable and I aim to show that it confronts Ted Sider's stage theory (1996, 2001) with a problem whose resolution is enlightening.

Imaginary fission scenarios have been presented in various ways such as divided brain hemisphere transplants and malfunctioning Star Trek teleporters. For the sake of conceptual clarity I shall imagine a case where a person's body is caused to divide smoothly in an amoeba-like way. This could be thought of as a complex process which involves the splitting of each constituent molecule into two molecules of just the same type followed by the spatial separation of the two sets of molecules. Such a process would require a huge input of energy, doubling the total corporeal mass involved, but fission of this type is not obviously physically impossible. Let's imagine that the subject is anaesthetised and that their body is put into a hi-tech "splitter". The two emergent bodies are then aroused in rooms of different colours, one green, the other blue.

According to Sider's stage theory what a person is at any time is a stage in their history. Stage theory embraces the idea of temporal parts but rejects conventional reductionism about identity, which identifies a person's body with an aggregate of temporal parts, a so-called world-tube or spacetime worm. This feat is achieved by exploiting the idea of counterparts which was used by David Lewis in his theory of possible worlds (1986). Temporal parts are treated in Sider's stage theory as a type of

counterpart. Thus I am now the same person as the boy who scrunped apples because I have past temporal counterparts who scrunped apples. My having those temporal counterparts makes it true that I was that boy just as, given Lewisian possible worlds, my having modal counterparts with blond hair makes it true that I might have been blond.

According to stage theory it is true that I will be bald iff I have future counterparts who are bald. If I'm worried about being bald then why, an objector to stage theory might ask, should I worry about having a bald future counterpart? After all, that counterpart is not me, he is someone else. The stage theorist might respond that I should be worried about my bald future counterparts because they are people who I will be. I shall be assuming that some such response is satisfactory but I am not here claiming that stage theory is the best analysis of transtemporal identity. My point is going to be that if we take stage theory seriously in contexts of personal fission then it seems to pose another problem which is worth addressing.

We shall be wanting a hypothetical subject who, for whatever reason, favours the stage-theoretic analysis. The Tuscan sculptor Stagio Stagi had the right name to be a stage theorist if anyone ever had so let's call our hypothetical subject Stagio. In matters of transtemporal identity Stagio embraces Sider's stage theory. According to stage theory, if Stagio is about to enter the splitter it is true that he will awake in the green room and that he will awake in the blue room (Sider 1996: sec.II; 2001: 201). That is because Stagio entering the splitter (a temporal stage) has a future counterpart in the green room and a future counterpart in the blue room and Stagio prior to fission has the relation *will be* to each of those counterparts. The fact that each of those simultaneous future counterparts is a distinct person is not at odds with this result. According to stage theory then, if Stagio enters the splitter he should expect to awake in the green room and he should expect to wake in the blue room.

What is Stagio to make of this ? If he is told that he will be imprisoned and tortured in the blue room but will be free to leave the green room how worried should he be? There is a state of mind with which we are familiar which *prima facie* seems suggestive of the attitude Stagio might to adopt here. It is probabilistic expectancy. If Stagio were told that rather than going through the splitter he would be tortured or not depending on the toss of a fair coin he would be pretty worried but he would prefer that option to certain torture. Ought Stagio's level of concern at the prospect of fission be the same as that in the corresponding chancy scenario? I shall argue that there is good reason to think so.

Return to the original fission scenario and call the person who awakes in the green room StagioG and the person who awakes in the blue room StagioB. These are distinct persons who were both previously Stagio. And imagine that when StagioG and StagioB are aroused the lights are out in their

rooms so that they are in complete darkness whilst both are fully informed about what the fission setup is¹. StagioG can very plausibly assign probabilities as to whether he is in the green room or the blue room. Given the symmetry of the setup it seems reasonable that he should assign equal probabilities to each possibility and so a probability of a half to each. StagioB should do likewise. So StagioG and StagioB should each think they have an evens chance of finding themselves in the green room or in the blue room when the lights are switched on.

Stagio valued greenbacks. He had greenbacks in the bank prior to fission and he could have been sure that after fission he would be in a situation where another person had equal claim on his bank account unless provision were made to set up bank accounts in the names of StagioG and StagioB which could each draw on as many greenbacks as were in Stagio's account prior to fission. To be fair to Stagio, let's suppose that such dual accounts have been set up.

Stagio was a betting man and so are StagioG and StagioB. Each is told that for a stake of one greenback they will receive three greenbacks if in the green room and nothing if in the blue room. If confident in the honesty of the setup both should accept this as a good bet since each thinks he has an evens chance of being in the green room and so both StagioG and StagioB should pay the stake.

Now imagine that Stagio is told in advance of fission what the payoff regime is going to be and is told that StagioG will only receive the payoff if the stake is paid before fission. Stagio knows that when the lights are off StagioG and StagioB will both think that the bet is worth taking and that each will regret that they had not paid the stake when they were Stagio if it is indeed not paid. That seems to be a good reason for Stagio to pay the stake in advance.

Go on to consider the scenario without blackout where Stagio is simply told that StagioG will receive thrice the stake paid prior to fission and StagioB will receive nothing. Stagio knows that StagioB will be disappointed to have lost but he also knows that StagioB will realise that if the lights had been out then he, StagioB, would have taken the bet. A person can regret having lost a bet but it is unreasonable to thereby regret having taken the bet if it was a good one. So counterfactual post-fission blackout is enough for Stagio to make a judgement about his future attitudes to the bet.

So if Stagio is told that he will undergo fission into StagioG and StagioB, that the bank account setup will be as described and that StagioG will receive a three-to-one payoff for a stake laid prior to fission then Stagio should leap at the chance if he's a gambling man. Which means that Stagio's betting behaviour prior to fission should be exactly what it would be if he were not faced with fission

¹ This idea is taken from Lev Vaidman (1998) who employs it in the context of the Everett interpretation of quantum mechanics.

but rather with a gamble on the tossing of a fair coin. That can seem strange because Stagio prior to fission can be sure that he will lose the bet. But of course he can also be sure that he will win the bet, which is no contradiction according to Sider's stage theory.

We have been imagining a symmetric bifurcation of Stagio's body but splitting into three, four or even a denumerable infinity of output bodies ought to be allowable if bifurcation is. If there were a number N of distinct outcomes then output persons held in a state of ignorance about their outcome should very plausibly assign a probability of $1/N$ to that outcome being a particular one of the range. This makes the relevant counterfactual future credences which guide Stagio's pre-fission betting unequivocal. For instance, if Stagio is faced with two possible fission pathways, one where he undergoes a three-way split and the other where he undergoes a double two-way split, with one output body from the first split, prior to arousal, undergoing another two-way split, there is no ambiguity about how Stagio should judge a bet. The fact that there will be three outcomes when he is aroused is all he needs to consider.

The upshot is that although Stagio must entertain a novel form of expectancy when faced with fission his most paradigmatically self-centred behaviour, betting for personal gain, ought rationally be identical to what his betting behaviour would be in a chance situation where the possible outcomes correspond to what are the actual outcomes post-fission and where the numerical proportion of fission output branches for a particular outcome is equal to the corresponding chance for that outcome in the equivalent probabilistic situation. This novel form of expectancy is unfamiliar but can be as firm a basis for rational provision for one's personal future as is probabilistic expectancy. It might be called paraprobabilistic expectancy, the prefix referring to the idea of parallel lives.

References

- Lewis, D. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Sider, T. 1996. All the world's a stage, *Australasian Journal of Philosophy* 74: 433-53.
- Sider, T. 2001. *Four Dimensionalism*. Oxford: Oxford University Press.
- Vaidman, L. 1998. On Schizophrenic Experiences of the Neutron or Why We Should Believe in the Many-Worlds Interpretation of Quantum Theory, *International Studies in the Philosophy of Science* 12 : 245. Available online at : <http://xxx.lanl.gov/abs/quant-ph/9609006>