

# Justifying Rational Choice: the role of success\*

Bruno Verbeek

Leiden University, Netherlands

## §1 Introduction

The theory of rational choice can be interpreted in several ways. One can regard the theory as a representing the choices of agents. The theory is interpreted as an empirical hypothesis for further research. Alternatively, one can regard the theory as an axiomatic modeling assumption for social theory. However, in this essay I will not discuss these descriptive and predictive interpretations of the theory. I will be concerned with the normative interpretation of the theory. On this interpretation the theory of rational choice is a systematic account of how agents ought to choose so as to realize their goals or preferences. The theory is, therefore, instrumentalist. The theory is neutral with regards to the goals or preferences of the agent. It takes these as a given input for its recommendations.

So far, I have been talking as if there is one unproblematic account of how to choose. However, as we shall see, that is not the case. There are several competing proposals for the rational procedure of choice. How do we determine which one is correct? Since the theory is supposed to be instrumentalist and neutral it is only natural to assume that the actions recommended by the rational choice procedure should be successful; successful, that is, in terms of the goals and preferences of the agent. If a procedure fails to produce successful choices, it cannot be the correct procedure of choice. And if a choice is successful then the procedure that recommends it is *ipso facto* rational. This gives us two related claims about the role of success in the justification of a choice procedure. First, success is necessary to establish the rational acceptability of a procedure of choice. Secondly, success is sufficient to establish the rationality of the proposed procedure of choice.

These two claims together form the doctrine of *pragmatic foundationalism*.<sup>1</sup> It has been advocated by authors such as David Gauthier and Edward McClennen.<sup>2</sup> Gauthier has attacked standard game theory and McClennen has criticized standard decision theory. They did this by showing that there are situations where agents do worse than they would

---

\* I would like to thank Michael Baurmann, Don Bruckner, Bruce Chapman, Gijs van Donselaar, Govert den Hartogh, Hartmut Kliemt, James MacAllister and Mike Ridge, as well as various audiences that were subjected to earlier versions of this paper. Special thanks and recognition go to Wlodek Rabinowicz who gave detailed comments on two separate occasions. I am immensely grateful for this opportunity to thank him and congratulate him on this occasion.

<sup>1</sup> This term is coined by Edward McClennen in (McClennen 1990, p. 4-5).

<sup>2</sup> (Gauthier 1986; 1994; 1997) and (McClennen 1990).

on their alternative theories (*constraint maximization* and *resolute choice* respectively), thus using success as a necessary condition. Moreover, in defending their own views they claim that since their alternative theory is more successful than the standard theory this shows the rational superiority of resolute choice and constraint maximization. Therefore, they treat success as a sufficient criterion for rational acceptability.

Appeals to claims of pragmatic foundationalism are found not only in the periphery of rational choice literature. For example, the first claim, i.e., that success is necessary, is used to argue against intransitive preferences. The *money pump* argument demonstrates that agents with such orderings fail to realize success in their own terms. Similarly, the *Dutch book* argument shows that agents whose probability assignments do not satisfy the standard rules of Bayesian probability calculus will fail to be successful.<sup>3</sup>

The claims of pragmatic foundationalism seem relatively unproblematic in the context of so-called normal form decision problems under certainty. In such contexts the agent has to make only one choice to realize the desired outcome, and chance does not play a role at all. In such a situation, if the choice procedure does not recommend the best or one of the best outcomes, surely it fails as a rational procedure. Similarly, if X is the best outcome, and there is a procedure that recommends X, that procedure must be a rational one.

Things become more complicated when we introduce probabilities and uncertainty. In such situations the relation between success and rationality is not as straightforward as it is under certainty. For example, suppose the agent faces a choice between \$1 for sure and a lottery that pays \$100 with probability .1. Assuming the agent cares only about money and has a neutral attitude towards risk (i.e., her utilities are a positive linear function of money) it is not straightforward which recommendation a rational procedure of choice would give. If the procedure recommends accepting the lottery the agent might end up with nothing. Is that sufficient to reject the procedure as a rational procedure? Many would be inclined to deny this. It would be a case of bad luck. Arguing that the recommendation is wrong if the lottery does not pay is committing the “bad-outcomes-bad-decision-fallacy”.<sup>4</sup> Therefore, in contexts where chance plays a significant role a beneficial outcome can be absent even though the action is rational. Success is not necessary to establish the rationality of the action or the procedure that recommends it.

On the other hand, suppose that the lottery does happen to pay. Arguing that the recommendation made by the choice procedure is rational in such cases is committing the conjoint “good-outcomes-good-decision-fallacy”. Therefore, the presence of a beneficial outcome is not sufficient to guarantee the rationality of the action in choice under uncertainty. Success is not sufficient to establish the rationality of the action or the procedure that recommends it.

---

<sup>3</sup> Whether these arguments do show what they are supposed to show is a much-debated matter. See (Hampton 1998) and (Schick 1986).

<sup>4</sup> (Frank 1988, p. 72-75)

This is, of course, just what expected utility theory would recommend. In choice under uncertainty it is not actual success that determines the rationality of a choice procedure but rather the expectation of success. This introduces the question which expectations matter and how the agent should deliberate with regards to them. The typical answer is that the expectations of the agent must comply with Bayesianism. This position has been criticized by Mark Machina and others who argue that certain assumptions of probability theory are not normatively secure.<sup>5</sup> In particular, the so-called *independence assumption* is called into question.

In this essay I will abstract from these problems of choice under uncertainty and concentrate instead on another context of choice, to wit, choosing over time. I will discuss only those decision problems over time which involve certainty. As we will see there are special problems for decision making over time which are absent from the one-shot case. These problems cast a new light on the idea that success is part and parcel of the justification of a choice procedure. I will argue against pragmatic foundationalism. Success is neither sufficient nor necessary to establish the rational acceptability of a choice procedure.

The remainder of this essay is organized as follows. Section two introduces three plausible requirements for choice over time. Section three introduces the problem of unstable preferences and argues that the requirements introduced in section two cannot all be met when the agent has unstable preferences. In addition, I give a pragmatic argument that uses success to identify the correct procedure of choice in those cases. Sections four and five discuss a general problem with this type of argument. The result of which is that success is not sufficient to establish the rational choice procedure and that there are reasons to doubt it is necessary. Section six strengthens this general conclusion. A proper understanding of the simple pragmatic argument presented in section four shows that success is not necessary to establish the rational acceptability of a choice procedure. This argument is only valid if we can demonstrate that there are rational plans in the absence of successful outcomes. Section eight provides such a demonstration. As it turns out, all three procedures can be rationally acceptable under circumstances. The context of the choice situation determines which procedure is acceptable and should be applied. I conclude that pragmatic success does play a role in the evaluation of the different procedures. However, this role cannot be characterized as a necessary or a sufficient condition. Section eight concludes with some remarks about the role of success in the justification of a choice procedure. I speculate that the best model to understand this role is not a foundationalist model of justification but a coherentist one.

---

<sup>5</sup> (Machina 1989).

## §2 Conditions of planning

Edward McClennen has formulated three intuitively plausible requirements of rational choice over time that enable us to characterize the different procedures of choosing over time available to the agent. These can be illustrated in figure 1. Suppose an agent faces the choice between three outcomes A, B, and C. However, these outcomes are not all directly available. That is, she could either choose B directly at  $t=1$ , or make another choice, to be faced with the choice between A and C at  $t=2$  (see figure 1).

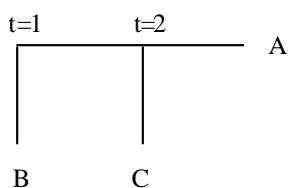


Figure 1

A *plan* is a detailed specification how to choose at each choice point that can be reached by the application of the plan. Thus, if a plan calls upon the agent to go “right” at  $t=1$  the plan will prescribe how she should continue at  $t=2$ . It seems plausible to require that a rationally acceptable plan remains acceptable during its execution. In other words, if a plan is acceptable at  $t=1$  its continuation at  $t=2$  should be acceptable as well. An acceptable plan consists of acceptable plan continuations. This is the first requirement of rational plan. A rational plan should satisfy *dynamic consistency* (DC).<sup>6</sup>

Secondly, the number of decisions or the sequence of the decisions should not determine whether a plan is acceptable or not. A rational agent should plan to realize the same outcome whether the situation is such that she has to make just one choice or several to execute the plan.<sup>7</sup> More precise, the rationally acceptable plans in the extensive-form and the normal form have the same outcome. This is the requirement of *normal and extensive form coincidence* (NEC).<sup>8</sup> Thus, a rational agent would select a plan in figure 1 if and only if she would plan for the same outcome in figure 2.

---

<sup>6</sup>(McClennen 1990, p. 120). Dynamic consistency is plausible only if there is no (unforeseen) change of information between  $t=1$  and  $t=2$ . If there is such a change, a plan continuation that seemed acceptable at  $t=1$  might no longer be acceptable because the agent would not have adopted such a plan in the first place had she known what she knows now.

<sup>7</sup> This assumes that the process of decision making itself does not alter the (value of the) outcomes in any way. This need not be the case. However, it seems a plausible assumption in the “small” choice problem we are considering in figures 1 and 2.

<sup>8</sup>(McClennen 1990, p. 115).

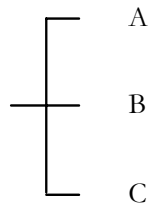


Figure 2

Since an acceptable plan is supposed to consist of acceptable plan continuations we need to determine what counts as an acceptable continuation. A rational agent is concerned about the outcomes of her choices. Therefore, each plan continuation should be consistent with this forward-looking nature of deliberation. This brings us to the third and last of the requirements, the requirement of *separability* (SEP).<sup>9</sup> A continuation is acceptable if and only if it is acceptable as a brand new plan from that choice point onward. In other words, the acceptable plan continuation in figure 1 at t=2 should correspond to the acceptable plan in the situation where the agent where to find herself facing just A and C.

### §3 A pragmatic argument for resolute choice

Standard expected utility theory assumes that the preferences of the agent are consistent. One of the conditions of consistency is the so-called *alpha* condition.<sup>10</sup> It requires that the preference order remains stable when the set of available outcomes becomes smaller. More precisely, if the agent prefers A to B when the available set of outcomes contains A, B and C, a rational agent is expected to prefer A to B once C is no longer available.

More often than not one needs to make a series of choices before realizing the preferred outcome. In the course of such a series the set of available outcomes usually becomes smaller. What was an available outcome at t=1 need no longer be available at t=2 because of the choice(s) one has made. Suppose that my preference ordering over outcomes satisfies all the requirements of expected utility theory, including the requirement that it is stable. In that case I can simply choose at each time the most preferred outcome among the outcomes still available while being sure that I will end up with my ex ante most preferred outcome. Following my present preference guarantees an optimal choice in the sense that classic decision theory recommends.

---

<sup>9</sup>(McClennen 1990, p. 122).

<sup>10</sup>(Sen 1970). Sen formulated alpha in the context of choice functions rather than preference orderings. In this essay will ignore this, since the revealed preference theorems suggest that we can reconstruct the preference order of the agent by investigating the range of the agent's choice function for each subset of the set of options. The alpha condition allows the theorist to reconstruct the preference ordering of the agent by a complete pair-wise comparison of all available outcomes.

What if preferences are not stable? In such cases it is possible that following one's present preference leads to sub-optimal outcomes. Classic decision theory cannot tell us anything except that one should avoid having such preferences.<sup>11</sup> Many authors have argued that in circumstances such as these a rational agent should *plan* her choices.<sup>12</sup> Rather than simply following her preferences at each and every point in time a choice needs to be made, a rational agent should carefully reflect in advance about her course of action and plan how she will choose should she reach a certain decision. This raises the question how the agent ought to plan. In other words, we need to know the rationally required planning procedure.<sup>13</sup>

Using the three conditions identified in section two, we can characterize alternative planning procedures. If and only if the agent's preferences satisfy the requirements of classic decision theory, including the requirement of stability, her plan will satisfy NEC, DC and SEP.<sup>14</sup> Therefore, if an agent's preferences are not stable, it is to be expected that not all conditions can be met. Indeed this is the case. Suppose that an agent finds herself in the situation of figure 1. Suppose that her preferences are unstable. If the set of outcomes consists of A, B, and C she chooses A. However, she selects C if the options are limited to just A and C. If she were to face all three alternatives in the normal form (as in figure 2) she would choose the plan leading to A. Since NEC requires that her choice in figure 2 corresponds to how she would choose in figure 1 she should adopt the plan to go "right" at  $t=1$ , planning to go "right" again at  $t=2$ .

However, at  $t=2$  the set of available options for her is reduced to just A and C. Separability requires that she selects the plan continuation that leads to C, instead of A. Therefore, her planning does not conform to dynamic consistency. Her unstable preferences lead her to accept a plan that contains an unacceptable continuation. Such plans are *myopic*.<sup>15</sup> Most authors agree that myopic planning is a bad idea.<sup>16</sup> The main

---

<sup>11</sup> Apart from the fact that this piece of advice seems out of place in a theory that purports to be neutral with regard to one's preferences, it assumes that agents are responsible for their preferences in a way that is not entirely plausible. Typically, we find ourselves with the values and preferences that we have. Sometimes we may lament the particular preferences we have, and believe that there are reasons for wanting to have other preferences. However, it is unclear if such beliefs are always sufficient reason to change our preferences.

<sup>12</sup> For example, (Rabinowicz 1995). These might not be the only circumstances in which a rational agent ought to for a plan. (Bratman 1987) has argued forcefully that constraints on information and information processing as well as instances of indifference all warrant the formation of a plan. In this paper I abstract from these complications.

<sup>13</sup> There are clear links between these discussions in decision theory and recent developments in action theory, especially (Bratman 1987), who analyzes intentions as (part of) plans for future action. I will not go into these connections in this essay. However, I believe this is one of the most promising developments for the integration of the two dominant (types of) theories of human action, i.e., rational choice theory and action theory.

<sup>14</sup>For a formal proof, see (McClennen 1990, p. 129).

<sup>15</sup> The phrase is that of (Strotz 1956). David Gauthier pointed out to me that one could question whether myopia should be characterized as a planning procedure at all since the concept of planning seems to imply a commitment to its execution. A myopic agent, as she is characterized here, does plan in the sense of deliberating and accepting a plan of action. However, it is characteristic of myopia that one need not follow through with a plan even if there is no change of information during the execution of the plan. I decided to

reason is that myopia can result in self-defeating choices. We can see why this is the case in figure 1. Myopia makes that the agent ends up with C, the worst option. Therefore, there seem to be good pragmatic reasons to avoid myopia.

Several authors have argued that instead of being myopic, the agent should be *sophisticated* in her plans.<sup>17</sup> That is, the agent in figure 1 should anticipate that at  $t=2$  she no longer will opt for A but choose C instead. Therefore, so the argument goes, A is not a feasible plan, or as Wlodek Rabinowicz puts it, A is not *performable*. A plan is performable if the agent does not have a reason to deviate from the plan after she has started its execution.<sup>18</sup> In figure 1 the agent has only two performable plans. The plan leading to B and the plan leading to C. Since at  $t=1$  B is preferred to C the only acceptable plan is the plan leading to B. This way of planning satisfies DC (be it trivially as there is no separate continuation after the first move). However, it violates NEC, because in figure 1 this agent would select the plan leading to B, whereas in figure 2 she would select A. Therefore, pragmatic considerations seem to counsel sophistication in situations such as these. Sophisticated planning gives the agent B, her second best option overall but the best performable option under the circumstances.

However, we have to ask the question why the plan leading to A is not performable. It seems that this is due to a commitment to separability. Separability requires the agent to consider her options at each choice point as if she faces them for the very first time. Were the agent to ignore separability, the plan leading to A becomes again performable. Since A is the best option, pragmatic considerations seem to indicate that SEP should be rejected as a requirement of planning. The agent should plan to choose A and *resolutely* pursue that option at  $t=2$ . Notice that a resolute agent seems to do better than the sophisticated agent in contexts such as these. Whereas the best the latter can do is B, only the former can realize the optimal outcome A.

It seems then that success completely determines the rationally acceptable planning procedure. It is both necessary and sufficient to identify the rational planning procedure. An agent with non-standard preferences (in this case, preferences that violate alpha) should plan resolutely. We have a completely *pragmatic justification* for resoluteness.

---

present myopia as a planning procedure here because, as I will argue below, there could be circumstances where myopia is the most rational way of going about things even if the agent has unstable preferences.

<sup>16</sup> To name but a few: (Strotz 1956), (McClennen 1990) and (Rabinowicz 1995).

<sup>17</sup> For example, (Strotz 1956), (Elster 1979) and (Levi 1992).

<sup>18</sup> (Rabinowicz 1995). The notion of performability is closely related to the idea of backward induction. Indeed, the whole idea of sophisticated choice can be regarded as an implication of backward induction.

#### §4 Two conceptions of justification

However, for many – including me – this conclusion is too fast. In this section and the next, I will discuss an important argument against this conclusion. David Velleman has argued that success is neither necessary nor sufficient for determining the rationality of a principle of rational choice.<sup>19</sup> Whereas success arguably is an appropriate requirement of the choice of an action (or in our case that of a plan) it is not a proper determinant for the choice *how* to act or plan. This latter choice, the choice for a procedure that identifies the rational choice or the rational plan, cannot be guided by considerations of success. This latter choice is not the object of practical reasoning but of theoretical reasoning. Therefore, success is the wrong sort of criterion for assessing the correctness of this choice.

For suppose it were. That is, suppose that the correct procedure of rational planning is not something we *discover*, but something that is object of *practical deliberation* much in the same way as we deliberate about the choice of plan. That would mean that the evaluation of a proposed procedure of rational choice (i.e., the correct procedure for rational planning) is itself an instance of rational choice, which is supposedly constrained by the same procedure. This, so Velleman argues, begs the question of the rationality of that procedure. Consequently, demonstrating that the adoption of a particular procedure of rational planning (whether it is myopia, sophistication or resoluteness) brings success is irrelevant for establishing the rationality of that procedure itself.

I hesitate to endorse Velleman's conclusions. I share his intuition that the problem of identifying the correct procedure of rational planning is a matter of theoretical reasoning and not a practical choice. However, the claim that a pragmatic justification is question begging is acceptable only if one shares that intuition. Let me explain. Let us assume that the argument in favor of resoluteness is valid. That is, for pragmatic reasons one should choose a resolute planning procedure. Why would one accept this as an argument in favor of resoluteness? The answer the pragmatist gives us is that acceptance brings success. Suppose a critic would not be satisfied and would demand why success is a proper criterion for acceptance of resoluteness. The pragmatist cannot and would not give any other answer: acceptance brings success – period. This is question begging only if one thinks that the acceptance of an argument is a matter of belief, that is, if one thinks that such acceptance is a matter of truth. And this is exactly what the pragmatist will deny in this context. For her the proper ground for acceptance of a choice procedure is not whether it is appropriate but whether acceptance will bring success.<sup>20</sup>

---

<sup>19</sup> (Velleman 1997).

<sup>20</sup> Note that stronger pragmatic positions are possible as well. One can imagine a pragmatist who thinks of every instance of theory acceptance as a matter of choice which is to be determined by whether or not acceptance will bring success. Here I have limited pragmatism to the domain of theories of rational choice. An interesting further question, which I will not pursue here, is whether the classic American pragmatists were pragmatists in this sense.



It may seem that there is no difference between a procedure bringing success or it being appropriate given the background of instrumental rationality. However, things start to look very different if we look at situations where the procedure for deciding itself has consequences other than the choice it recommends. A good example is the so-called *toxin puzzle*.<sup>21</sup> The story is that an eccentric millionaire will give you a million dollars if you can form the plan now to drink a vial of toxin tomorrow that will make you sick for a day or two. Given that you prefer one million dollar to a few days of being sick should you be resolute and drink the toxin tomorrow? Tomorrow there is no reason to drink the toxin so a sophisticated chooser will not plan to drink it. A myopic chooser might form the plan but the millionaire will recognize that the plans of a myopic chooser are not worth the paper they are written on. Only a resolute chooser will be believed. Thus resoluteness pays here. However, it pays not (only) because resoluteness recommends the best plans but because the procedure itself is beneficial. This effect is relevant for its justification according to the pragmatist. His critic, however, will argue that this it amounts to arguing that if it “pays” to believe that resoluteness is rational you should believe it. It seems that this is not the right sort of criterion for belief acceptance.

What emerges here is that the pragmatist has an alternative picture of what a successful justification of a rational planing procedure should look like. Whereas the critic of the pragmatist position will accept a planning procedure if she believes such a procedure to be appropriate, the pragmatist will accept such a procedure only if doing so will bring success. How should we decide between these two rival conceptions of justification?

## **§5 Rational choice procedures: imperfect or pure?**

The different picture of justification is not the only thing that divides the pragmatist and her critic. There is a further difference in how they regard the status of the justified procedure of planing. This becomes apparent once we look at some of the characteristics of pragmatic justifications. First, any argument that demonstrates that the proposed principle in question systematically leads to sub-optimal results provides sufficient grounds for the pragmatist to reject that principle. Therefore, a successful pragmatic justification should be *self-supporting*. The application of a justified planning procedure should not have results that undermine the reasons for accepting it in the first place. A good example of failure on this count is the argument against myopia in the previous section.

The second characteristic is that self-support is extended to the acceptance of the theory. Any argument that demonstrates that the acceptance of a principle of choice itself is unsuccessful would count against that principle from the point of view of the pragmatist.

---

<sup>21</sup> (Kavka 1983).

Whereas the first type of consideration operates at the level of the *application* of the proposed procedure of choice, this consideration operates at the level of *acceptance* of the proposed procedure of choice. This characteristic is what sets pragmatic justifications apart from their non-pragmatic alternatives. An example of failure of this type is the standard objection against utilitarianism that acceptance of it actually leads to less overall happiness.<sup>22</sup>

In short, a successful pragmatic justification for a planning procedure needs to demonstrate that both the application and the acceptance of the theory will lead to success. This is all that such a justification needs to establish. There are no further questions as to why success would count as the proper criterion.

Suppose that there is at most one rationally acceptable planning procedure that satisfies this criterion.<sup>23</sup> If this is the case we can characterize such a procedure as a *pure* procedure of rational planning: following the procedure is both necessary and sufficient for realizing success. The procedure is sufficient for success since it is established by the pragmatic demonstration that its application leads to success. The procedure is necessary for success because it is the only procedure that could lead to success.<sup>24</sup>

The non-pragmatist has a completely different idea about the status of the justified planning procedure. A critic of pragmatic justifications like Velleman assumes a parallel between action and belief. To believe X implies that one believes X is true. However, whether it is rational to believe X depends on the procedures through which one came to the conviction that X. Although such procedures aim for the truth they are typically fallible in that respect. Therefore, the procedure for belief acceptance is an *imperfect* procedure. Following the procedure is neither necessary nor sufficient for the belief to be true. Similarly, the fact that X happens to be true is neither necessary nor sufficient for the rationality of the belief that X. One may come to believe X, a true belief, because of a completely spurious procedure.

Velleman's argument presupposes that the relation between the rationality of an action and that action having success is analogous to the relation between rational belief and truth. Whether or not a plan is rational depends on the procedure. However, whether or not the procedure is rationally acceptable is not determined by the success of the plans it

---

<sup>22</sup> There are ways for the utilitarian to avoid this criticism. For example, see (Pettit 1991).

<sup>23</sup> I will come back to the plausibility of this assumption in section 7. (Velleman 1997) argues that, unless one has specified exactly what "success" is, there are numerous candidates for principles or procedures which are self-supporting in this way. Almost any principle can provide the required self-support. For example, a principle which calls for unconditional cooperation in a prisoners' dilemma will be self-supporting if "success" is defined as "an effort at cooperation". Both the application and the acceptance of this principle will lead to efforts to cooperate. It might be objected that "an effort at cooperation" is an implausible notion of success. However, strictly speaking this type of objection is not open for the pragmatist since the fact that it does provide a self-supporting principle of choice is sufficient to establish the rationality of that principle of choice.

<sup>24</sup> Note that if there are more pragmatically justifiable procedures there is no pure procedure of rational planning but a set of *perfect* procedures.

recommends. If this is a correct way of thinking about the relation between rationality and pragmatic success, we have reasons to doubt that pragmatic success is necessary (let alone sufficient) to establish the rationality of a particular planning procedure. Just as a belief can be rational without being true, a plan can be rational without bringing pragmatic success.

So we do not just have rival conceptions of justification, we also have different pictures of the status of justifiable procedures of planning. Whereas the pragmatist thinks of these latter as pure procedures her critic thinks of them as imperfect procedures. These differences are related. If one has the intuition that success is the proper criterion for acceptance of a planning procedure, then one will characterize such a procedure as a pure procedure. This gives us a way to settle one issue between the pragmatist and non-pragmatist. Above I suggested that it is just an intuition whether or not success is the proper criterion for theory acceptance. We now have the tools to throw some light on this issue. For if it is true that belief and action are analogous, the critic has an important argument for the claim those rational planning procedures are imperfect procedures. If that is correct, then (*modus tollens*) it cannot be the case that the proper criterion for acceptance of a planning procedure is the question whether or not doing so will have success. Therefore, the question to answer is whether there is such an analogy between belief and action as Velleman has suggested.

As we saw above pragmatists reject this analogy between belief and action. David Gauthier argues for this rejection as follows:

A person's life may go better if he forms a belief that is not well supported by procedures directed at truth, and he may sometimes be in a position to recognize this. Although life may go better if he performs an action that is not well supported by the procedures directed at success, he cannot be in a position to recognize this at the time of performance and so cannot suppose it rational to eschew such procedures on that account. (Gauthier 1994, p. 700).

Stated in this way there is a disanalogy between the rationality of belief and that of action. For example, life in Russia under Stalin may go better for a person if she were to believe everything the communist party claims, even if this contradicts the outcome of procedures directed at the truth. Her life may go better because she will not be suspected by the security forces as a contra-revolutionary element.

However, the same is not true for action. A person might perform an action that is irrational although her life will actually go better as a result of it. For example, a person's life might go better if she were to buy a ticket in the national lottery in which the expected benefits are marginal in comparison to the cost of the ticket, if it turns out that it is in fact a winning ticket. However, she cannot be in a position where she realizes that this is the case (i.e., that it is in fact a winning ticket) and it not being rational to buy the ticket.

Therefore, rationality in belief and rationality in action do not stand in the same relation to one's life going better. Gauthier concludes that the analogy between the rationality of belief and that of action does not hold. Consequently, there is no reason to suppose that rational planning procedures are imperfect procedures.

I am not convinced that this argument disproves the analogy between action and belief. It does not state the analogy correctly. The question should be whether success (i.e., one's life going better) plays the same role for the rationality of an action, as truth for the rationality of a belief, rather than success. In the first line of the passage quoted above Gauthier talks about a person's life going better if he were to form an irrational belief. Further down he compares this with a person's life going better if he were to perform an irrational action. However, that is not the proper analogy. What should be compared to life going better is the totality of one's true beliefs. Thus, we should compare whether one can entertain more true beliefs if one ignores the procedures directed at truth with the question whether one's life could go better if one ignores the procedures directed at pragmatic success.

Once we state the analogy this way, there is a clear parallel between action and belief. It might be the case that one will entertain more true beliefs if one ignores the procedures directed at the truth in some particular case, but, just as is the case with action, one cannot be in a position to recognize this when one forms the belief. One cannot because such recognition will, in any plausible procedure for belief acceptance, play a deciding role. For example, it may be that ignoring the available scientific evidence for the relation between mass and the gravitational acceleration of the Earth will lead Galileo to come to believe a true belief, for example, that this acceleration is constant. However, Galileo cannot be in a position where he realizes that the gravitational acceleration is constant (e.g., through observation after dropping his linked cannon balls) and eschew the scientific *method* as a result because the scientific method will endorse his conclusions after this realization. The parallel between action and belief is not threatened by Gauthier's point as long as we assume plausible procedures for belief acceptance. Therefore, we have every reason to assume that the rationally justified planning procedure is an imperfect procedure. Success is neither necessary nor sufficient for establishing the rationality of a planning procedure. The pragmatist conception of justification is incorrect unless there is a different argument to the effect that rational choice procedures are pure. I am not aware that such an argument exists.

However, Gauthier's argument suggests a way in which success is relevant. A belief can be rational even if it happens to be false. Similarly, an action may be rational, even if it turns out to be pragmatically unsuccessful. Theoretical rationality and practical rationality are both imperfect procedures. However, one cannot, or rather, should not, continue to believe X, or believe it is rational to believe X, when one realizes that X is false. Similarly with action: one should not perform an action, nor believe the action is rational, when one

realizes it leads to failure. Moreover, this should be part of the procedure of deliberation. Therefore, success itself is not a necessary condition for the rationality of a rational planning procedure. However, the (justified) belief in success is a necessary condition for the rationality of a theory of rational planning.

If this is correct, pragmatic justifications are indeterminate. Following the rationally superior procedure is neither necessary nor sufficient for attaining success. Establishing that following a certain procedure leads one inevitably to success is not enough to demonstrate the rational acceptability of that procedure even if all alternatives are less successful. We need additional arguments to demonstrate the rational superiority of resoluteness or indeed any of the other planning procedures.

## §6 The inapplicability of pragmatic arguments to unstable preferences

In order to assess the plausibility of the complex, general and abstract observations of the last two paragraphs I propose that we look into the argument of section 3 again. There it was claimed that in the example of figure 1 the agent with unstable preferences realizes the best overall outcome (A) by being resolute. However, this conclusion is too fast. What warrants the assumption that A is the best outcome? Given the preference ordering of the agent we know that of all the three alternatives A is the best. However, once we restrict the domain of outcomes from {A, B, C} to {A, C}, A no longer is the preferred option. This is *all* we know. It is insufficient proof for the claim that A is the best outcome.<sup>25</sup>

There are three reasons why A may not be the best outcome. First, it could be the case that this agent's preferences are intransitive. That is,  $A > B$  and  $B > C$  but  $C > A$ . Note that this particular ordering of the alternatives is compatible with the information we have about the agent's preferences. If her preferences are intransitive there is no best outcome because there is always another outcome that is better in a pair-wise comparison. In other words, in this situation there is no best outcome. If there is no best outcome pragmatic arguments are inconclusive because they simply do not apply.

Secondly, we could doubt that A is the best outcome without it being the case that the agent has intransitive judgements as to what is best for her. Strictly speaking, transitivity requires that if  $A > B$  and  $B > C$  then  $A > C$ . One can satisfy this conditional while denying the antecedent. Thus it may be the case that A and B are incommensurable when

---

<sup>25</sup> In this connection it should be noted that the main advocate of resolute choice, Edward McClennen, would not endorse the conclusion that A is the best outcome in our example. McClennen claims that an outcome is best if and only if the ex ante self and the post ante self can agree as to what is best. In those cases, and only in those cases, one can claim the superiority of resolute choice. I disagree with McClennen that these are the only contexts of choice where resoluteness is rational. In the next section I give an example of the rationality of resoluteness in the absence of a best outcome. Furthermore, I am not convinced that in all cases where there is a clear best outcome resoluteness is required.

compared as a pair. A preference ordering of this type does not violate transitivity. However, it is not complete. Again, we have reason to doubt that there is a straightforward best solution even though this agent could select the preferable outcome if she were to face all three of them at the same time.

Finally, it could be the case that the ordering is complete (all alternative outcomes are commensurable) and that it satisfies transitivity whereas it still is unstable. That is, it could be the case that  $A > B$  and  $B > C$  and  $A > C$  but that  $C$  is the preferred outcome when all three are considered. In such a case the manner of valuation does not generate an unambiguous best outcome if we just look at the preference ordering.<sup>26</sup>

It seems then that we can dismiss the pragmatic argument for resoluteness in our example. We have formal reasons to doubt its applicability because we cannot be sure that there is a best outcome in the first place. The presence of such an outcome is essential for establishing the rationality of one of the three planning procedures if pragmatic considerations are necessary to establish the superiority of any of the procedures. If this is correct, we may have identified a class of examples of unstable preference orderings in which success is not even a necessary condition for the rationally acceptable planning procedure.

This tentative conclusion is only valid if there is at least one rationally superior planning procedure in cases like this. I believe we can identify such a planning procedure here. Which procedure is rational depends on the complete, “thick” description of the situation. As it turns out, pragmatic arguments do play a role in the identification of this procedure. However, this role is one that does not allow us to claim necessity (let alone sufficiency) for success.

## §7 Three cases of unstable preferences

How could such a “thick” description of the outcomes and the choice situation determine the rationally superior planning procedure when strict pragmatic considerations cannot? I will give three examples which each point to a different procedure.

Example 1: *the minimax regret chooser*.<sup>27</sup> Consider an agent who orders her prospects so as to minimize her possible regret. She considers for each prospect what could have happened under the same conditioning event had she chosen otherwise. Suppose this person faces a decision tree like that of figure 1. Suppose moreover that outcome  $A$  is a lottery which, depending on certain events ( $E_1$ ,  $E_2$  or  $E_3$ ), will either give her \$10, \$0 or \$3.  $B$  stands for a lottery that gives \$2, \$4 or \$10 under the same conditioning events.  $C$ ,

---

<sup>26</sup> See also the discussion in (Hampton 1998, p. 260-261, 276-278).

<sup>27</sup> (Savage 1972, ch. 9).

finally, will give \$10, \$5 or \$1 under those events. Suppose this agent lacks all knowledge of the likelihood of any of these states. In such a case she decides to take that course of action that will minimize her maximal regret. In order to determine this she looks at what she could have had under the conditioning event had she chosen otherwise. The maximum difference between what she actually got and what she could have had given the conditioning event is the amount of regret of that particular prize. She does this for each prize and each event and then she determines the maximum possible regret. Next, she opts for the lottery with the smallest maximal regret.

If we compare A, B and C in a table we can calculate the maximum possible regret (see figure 3). We see that the pattern of figure repeats itself here. When comparing A, B, and C, A is the most preferred option. However, if we limit the range of comparison to just A and C, as will be the case at  $t=2$ , C is preferred over A (see figure 4).

	E <sub>1</sub>	E <sub>2</sub>	E <sub>3</sub>	regret	E <sub>1</sub>	E <sub>2</sub>	E <sub>3</sub>	Max. regret
A	\$5	\$2	\$10	A	5	8	0	8
B	\$10	\$5	\$1	B	0	5	9	9
C	\$0	\$10	\$4	C	10	0	6	10

Figure 3, calculating the maximum possible regret when comparing A, B and C.

	E <sub>1</sub>	E <sub>2</sub>	E <sub>3</sub>	regret	E <sub>1</sub>	E <sub>2</sub>	E <sub>3</sub>	Max. regret
A	\$5	\$2	\$10	A	5	8	0	8
C	\$0	\$10	\$4	B	5	0	6	6

Figure 4, calculating the maximum possible regret when comparing A and C

In section six I argued that there is insufficient reason to assume that there is a best outcome in the example of section three. This could be due to intransitivity or because the ordering is incomplete. On the other hand, it could be the case that the preference order is unstable without any intransitivity or incompleteness. This first analysis is the correct one in the present example. This agent's preferences are intransitive..<sup>28</sup>

So how should we decide which is overall the rational plan? Inspection of the preference ordering does not help us. I propose we include in our considerations the standard of evaluation that generates the preference ordering, i.e., the avoidance of regret. One regrets one's choices when one realizes that one could have done better if one had

---

<sup>28</sup>An example of an alpha violation that does not imply intransitivity is the following. Let A have as prizes \$10, \$5 and \$6; B has \$0, \$10 and \$10; C, finally, has \$7, \$2 and \$4. Now C comes out best when comparing all three lotteries, whereas  $A > B$ ,  $B > C$  and  $A > C$  and transitivity is, therefore, maintained.

chosen differently given what happened. Regret is a holistic way of evaluating outcomes. Each outcome is judged against the background of what could have been the case had the agent chosen otherwise. This means that the agent should include outcomes that are no longer available as result of previous choices in the assessment the value of the outcomes. In other words, at  $t=2$  she might still regret not having chosen “down” in figure 1 to realize B. Even if she does not, as is the case in this example, she should still include B in the assessment of the choice ahead of her at  $t=2$ . Therefore, given her standard of evaluation she should plan to realize A and stick to that plan. Resoluteness then is the most rational way of planning in this case.

Several things should be noted about this conclusion. First, we arrived at this conclusion only after closer inspection of the situation. The preference ordering alone did not supply us with enough information about the best way of planning. Secondly, success did play some role in our conclusion. After we established the true nature of the standard of evaluation of the minimax regret chooser we could point to A as the best outcome and ignore the fact that B comes out best in all pair-wise comparisons. It is clear that success is not sufficient to single out resoluteness as the rational planning procedure. Is it necessary here? At this point the example is unclear because the “successful” character of A depends on the thick description of the standard of evaluation. Therefore, we cannot assess how this argument fares if we keep its complete description intact minus the pragmatic successful character of A.<sup>29</sup>

Example 2: *The potential addict*. Suppose that in figure 1 going “up” refers to partaking of a highly addictive substance, which will give a very pleasurable but fleeting sensation. Going “down” refers to abstinence. Outcome A should be interpreted as “have some fun, stop taking the drug and stay healthy”. B is “continue taking the drug and become addicted” whereas C is “have no fun but stay healthy”. Judged overall outcome A is preferable. However, once the potential addict has gone “up” he foresees he will have developed such a craving for continued drug use, he will no longer prefer A to C, but C to A. So we have the preference pattern of the example of figure 1.

Does this mean that we also have reason to endorse the conclusion that the potential addict should be resolute? I doubt that even the staunchest defender of resoluteness would be willing to defend that conclusion. There are two reasons for this. On the one hand our experience with addictions tells us that a resolute counter-preferential choice at  $t=2$  is not feasible. I mean this in a strong sense: it requires too much will power of the agent to

---

<sup>29</sup> One may be tempted to argue that this closer inspection of the thick description of the situation, resulted in a re-description of the outcomes, such that the preference order no longer is unstable and that, therefore, there is a clear best outcome available. That would strengthen the case for the necessity, if not sufficiency, of pragmatic arguments. However, one should resist such re-descriptions for reasons best explained by (Hampton 1998, 268-281). See also (Verbeek 2001).



execute the resolute plan.<sup>30</sup> Note that the doubt of the feasibility of A is not due to a normative commitment to SEP but has to do with other reasons.

On the other hand we do assume that most addicts deplore their addiction. They wish they were not addicted. Therefore, we can assume that although the potential addict prefers C at  $t=2$ , he will not do so in hindsight.

Both these considerations point to sophistication as the rational procedure. The first consideration tells us that not partaking of the drugs is a good way of protecting us from the consequences of a threatening lack of will power. Precisely because the potential addict cannot “trust” her future self to choose A, she settles for B, which is the second best outcome ex ante. The second consideration is a reason to be sophisticated because that is the way to protect our long-term interests in a healthy, drug free life-style.

Here, as before, we see that we can identify the rational planning procedure only after a closer inspection of the situation and the alternatives. The preference ordering alone provides insufficient information. Pragmatic success plays a role in this argument. Since we cannot be sure that outcome A is feasible, we have a choice between B and C. Based on the second consideration about the long term evaluation of C we chose B since that is the best under the present circumstances. Does this mean that success is a necessary condition here? Again I hesitate. We cannot isolate the relative success of B from the example, in order to test this hypothesis.

Example 3: *Mark and the girls*.<sup>31</sup> Mark is a pre-teen boy who firmly believes that girls are totally yucky. However, he also observes that boys slightly older than he suddenly become gaga over girls. Although he fails to see why they have such unstable preferences, he fully expects that he will experience the same type of preference shift. Luckily (...), there is an alternative. He could enroll in an all-male military boarding school and thus avoid being around girls.<sup>32</sup> Of course he would hate the school but it is a sacrifice he is willing to make to avoid the utterly foolish behavior he observes in others.

Note that Mark’s predicament corresponds to the preference shifts in figure 1. At  $t=1$  he prefers A, staying in his neighborhood school and not associate with the yucky creatures. Second best, from the point of view at  $t=1$ , is enrolling in the boarding school (B). However, at  $t=2$ , which is actually a year or so later, he strongly prefers to associate with girls (C) to A.

Should Mark be resolute and stick to his plan not to associate with these yucky girls? Does he have reason to believe he will not be able to live up to such a commitment and sophisticatedly enroll in the boarding school? The answer is, as I am sure anyone would

---

<sup>30</sup> Does this also mean that the potential addict at  $t=2$  ceases to be rational? That depends on the extent to which one is willing to include will power in one’s overall account of practical rationality.

<sup>31</sup> I borrow this example from (Gauthier 1997).

<sup>32</sup> Mark could also opt for a place in a Catholic seminary; another all-male environment. Needless to say that both seminaries and all-male military boarding schools have certain disadvantages.

agree, that he should do neither. Mark should be myopic in this case. Two considerations could be invoked to defend this obvious conclusion. One could claim that Mark's present preferences are irrational since they are ill informed. Mark fails to appreciate the value of a life in which there is place for loving and sexual relations with others. The sometimes embarrassing behavior teenagers go through is a necessary step in the realization of this value. Thus, Mark should ignore his present preferences and do nothing. "Going gaga over girls" is the best thing that could happen to him. Alternatively, one could argue that the anticipated switch in his preferences reflects a change in his values (or rather, in his standards of evaluation).<sup>33</sup> This is nothing to be afraid of. It comes with becoming a mature person that one's values change. Planning in order to avoid the consequences of such changing values reflects an irrational fear for a perfectly natural development on one's life.

Again, we see that there is a superior planning procedure although we cannot infer this from the pattern of preferences. In fact, here we have a case where we tend to ignore Mark's actual preferences altogether and appeal to considerations of value. Note that both responses accuse Mark's present preferences of irrationality. As such, they go beyond the realm of instrumental rationality. Finally, note that – again – pragmatic considerations do play a role in this example. Both arguments are to the effect that B, going gaga over girls, is in fact the best outcome for Mark. This does not give us any reason to suppose these considerations are necessary to justify myopia in cases such as this. The considerations we invoked to defend the rationality of myopia here are such that they automatically imply some sort of judgment as to what is the best outcome in this case. Thus, again we have no way to test the necessity of there being a best outcome for the identification of the rational planning procedure.

We can conclude the following. Success is not sufficient in any of the three cases to determine the rationally acceptable way of planning since we need additional arguments for this determination. Secondly, pragmatic considerations do play a role but we are unable to assess whether this role is a necessary one. Third, in so far as success is part of the argument in favor of a planning procedure in these three examples it is compatible with all three procedures (including myopia) identified in section 3. Therefore, pragmatic justifications (if indeed we can characterize the discussion in these examples as such) are completely indeterminate. Finally, success does play some role in all the examples. However, we were unable to test the necessity of success. This raises the question how we should understand its role.

---

<sup>33</sup> This sets this example apart from the example of the minimax regret chooser. The latter maintained the same values (even though her preferences shifted), whereas Mark's values change

## §7 The role of pragmatic success in defending a conception of rationality

In the previous sections I argued against the view that pragmatic considerations are both necessary and sufficient for justifying the rationality of a planning procedure. However, we also saw that such considerations are relevant in justifying the rational planning procedure. How should we think of the role of success in defending rational choice?

One proposal is that we can formulate a weaker criterion than success for justifying a planning procedure that is in the spirit of pragmatic foundationalism.<sup>34</sup> This condition would state that rather than success we should consider choice procedures in terms of non-failure. Non-failure could be seen on this proposal as a sufficient condition or a necessary condition. Thus, if a choice procedure does not fail to bring success then it must be justified (non-failure is sufficient). Vice versa, if a planning procedure is rational then it does not fail to bring success (non-failure is necessary). If this suggestion is going to work at all we should be careful and not identify non-failure with success. (For in that case the sufficiency of non-failure just is the sufficiency of success whereas the necessity of non-failure is the necessity of success). *Tertium datur*, therefore.

Let us start with the first suggestion. Non-failure could be sufficient for the rationality of a choice procedure. Since there is no formal characterization of success available in the cases above, none of the procedures fail. That would imply that all three procedures are appropriate in each case. Since I have demonstrated that this is not true, non-failure is not sufficient to establish the rationality of a choice procedure.

The second suggestion is that non-failure is necessary for the justification of a choice procedure. This seems right. It is applicable in all three cases that were discussed in section six. There we saw that all three choice procedures do not fail. We also saw that further considerations were needed to identify the appropriate choice procedure in each case. Perhaps this gives us the proper understanding of the role of success. It explains why non-failure alone does not tell us very much about the justified choice procedure. Moreover, it does capture part of the intuition that success is relevant in a rational choice procedure.

However, I believe this is still too strong. If our characterization of rational choice procedures as imperfect procedures is correct we cannot claim that non-failure is necessary for a procedure to be rational. We can expect that rational choice procedures sometimes will fail. The very most we can claim is that we should have the justified belief that our choice procedures will not fail in the richly described circumstances.

Can we say something more general about the justification a choice procedure other than that “it depends on the circumstances”? In what follows I can only briefly speculate about this and state my intuitions. Why would one believe that success is both necessary and sufficient for justifying a rational procedure of choice? I suspect that there are two

---

<sup>34</sup> This was suggested to me by Michael Ridge.

reasons for this conviction. First, there is the view that a successful justification must be a *foundationalist* one. Foundationalism in this connection is the view that a successful justification of a procedure of choice is based on a firm and unproblematic fundamental assumption. The foundation in this case is the consequentialist intuition that rationality makes one's life go better.

As we have seen, this is not a sufficient basis to justify a procedure of choice. Moreover, I suspect one will never be able to point out one alternative, absolute foundation that is sufficient to carry all the justificatory weight. For this reason I am more optimistic about the chances of a coherentist form of justification. In such a justification success is but one of the many considerations that can be introduced to support a particular procedure of rational planning.

The second reason why people might be inclined to think that pragmatic success is the bedrock of a theory of rational planning is the idea that there is at most *one* truly rational planning procedure. As the discussion of the examples in section seven has shown, this is not the case. If there is indeed only one rationally acceptable planning procedure it is a maximally permissive one that is compatible with any of the three procedures of planning depending on the situation.

Such a maximally permissive planning procedure again suggests a coherentist model rather than a foundationalist one because a coherentist account is more responsive to all the different and subtle considerations that could make a difference in specific situations.

Obviously this is not the occasion to spell out such a coherentist justification or to delineate the exact content of a maximally permissive theory of choice.<sup>35</sup> I hope to have shown, however, that we need other ideas than the notion of pragmatic success alone to pull off that project.

---

<sup>35</sup> (Radzik 1999) formulates an attractive coherentist model of justification of a theory of practical reason. As for the content of the desirable theory of choice, the proposal of (Rabinowicz 1995) for a theory of *wise choice* may be a good starting point since it rejects both NEC and SEP. However, more thinking is needed as to how wise choice would fare in the type of cases that I have described here, in particular, in the case of example 3.

## References

- Bratman, M. E. (1987). *Intention, Plans, and Practical Reason*. Cambridge, Harvard University Press.
- Elster, J. (1979). *Ulysses and the Sirens*. Cambridge, Cambridge University Press.
- Frank, R. (1988). *Passions within Reason*. London, W. W. Norton & Company, Inc.
- Gauthier, D. (1986). *Morals by Agreement*. Oxford, Clarendon Press.
- Gauthier, D. (1994). "Assure and Threaten." *Ethics* **104**: 690-721.
- Gauthier, D. (1997). "Resolute Choice and Rational Deliberation: A Critique and a Defense." *Noûs* **31**(1): 1-25.
- Hampton, J. (1998). *A Theory of Reasons*. New York, Cambridge.
- Kavka, G. (1983). "The Toxin Puzzle." *Analysis* **43**: 33-36.
- Levi, I. (1992). Feasibility. *Knowledge, Belief and Strategic Interaction*. C. Bicchieri and M. L. Dalla Chiera. Cambridge, Cambridge University Press: 1-20.
- Machina, M. (1989). "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty." *Journal of Economic Literature* **27**: 1622-1668.
- McClellan, E. F. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge, Cambridge University Press.
- Pettit, P. (1991). Consequentialism. *A Companion to Ethics*. Cambridge, Blackwell.
- Rabinowicz, W. (1995). "To Have One's Cake and Eat It Too: Sequential Choice and Expected-Utility Violations." *Journal of Philosophy* **92**(11): 586-620.
- Radzik, L. (1999). A Coherentist Theory of the Authority of Practical Reason.
- Savage, L. J. (1972). *The Foundations of Statistics*. New York, Dover.
- Schick, F. (1986). "Dutch Bookies and Money Pumps." *Journal of Philosophy* **83**: 112-119.
- Sen, A. (1970). *Collective Choice and Social Welfare*. San Francisco, Holden Day.
- Strotz, R. H. (1956). "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies* **23**: 165-180.
- Velleman, J. D. (1997). Deciding how to Decide. *Ethics and Practical Reason*. G. Culy and B. Gault. Oxford, Oxford University Press: 29-52.
- Verbeek, B. (2001). "Consequentialism, Rationality, and the Relevant Description of Outcomes." *Economics and Philosophy* **17**(2): 181-205.