Egocentric framing - one way people may fail in a switch dilemma: Evidence from excessive lane switching

Navon, David and Kaplan, Todd and Kasten, Ronen

University of Haifa, University of Haifa, University of Haifa

18 September 2013

Egocentric framing - one way people may fail in a switch dilemma: Evidence from excessive lane switching

David Navon Todd R. Kaplan Ronen Kasten
The University of Haifa, Haifa, Israel

Corresponding author:

David Navon
Department Of Psychology
The University of Haifa
Haifa, Israel 31905

Fax: 972 4 8240966
Tel: 972 4 8240927

Email: dnavon@psy.haifa.ac.il

Abstract

To study switching behavior, an experiment mimicking the state of a driver on the road was conducted. In each trial participants were given a chance to switch lanes. Despite the fact that lane switching had no sound rational basis, participants often switched lanes when the speed of driving in their lane on the previous trial was relatively slow. That tendency was discerned even when switching behavior had been sparsely reinforced, and was especially marked in almost a third of the participants, who manifested it consistently. The findings illustrate a type of behavior occuring in various contexts (e.g., stocks held in a portfolio, conduct pertinent for residual life expectancy, supermarket queues). We argue that this behavior may be due to a fallacy reminiscent of that arising in the well-known “envelopes problem”, in which each of two players holds a sum of money of which she knows nothing about except that it is either half or twice the amount held by the other player. Players may be paradoxically tempted to exchange assets, since an exchange fallaciously appears to always yield an expected value greater than whatever is regarded as the player’s present assets. We argue that the fallacy is due to *egocentrically* framing the problem as if the “amount *I* have” is definite, albeit unspecified, and show that framing the paradox *acentrically* instead eliminates the incentive to exchange assets. A possible psychological source for the human disposition to frame problems in a way that inflates expected gain is discussed. Finally, a heuristic meant to avert the source of the fallacy is proposed.

Keywords – Decision making; Reasoning; Cognitive fallacies and biases; Switching behavior

PsycINFO classification- 2340

Egocentric framing - one way people may fail in a switch dilemma: Evidence from excessive lane switching

1. Introduction

Much of human behavior tends to be stable, a pattern which researchers often attribute to risk aversion (see, e.g., Bell, 1982, 1988; Kahneman & Tversky, 1984) or to the status quo bias (e.g., Samuelson & Zeckhauser, 1988). In some cases, people seem reluctant to decide in favor of a switch of state even when doing so is likely to end in a considerable gain. It is not surprising that they behave so when the prospective gain is not intuitively transparent, as in the Monty Hall problem (see Granberg and Brown, 1995; Petrocelli & Harris, 2011), but sometimes they do that even when irrationality is reasonably salient, as when people continue to throw good money after bad in the face of sunk cost (see, e.g., Arkes and Blumer, 1985; Garland, 1990).

Nonetheless, people sometimes opt for the contrary - a switch of state or a change in behavior in the light of slim evidence that such a change might be beneficial. Some people do it often enough to deserve being called "switch seekers".

Switch seeking in the strong sense may be driven by exceptional discontent with what one has, by unrealistic expectations of what one could alternatively have (possibly induced by overevaluation of what others have), or by idiosyncratically high attraction to the change itself. Switch seeking of that sort seem confined to a small minority of persons, presumably having distinct personality characteristics.

Could it be, though, that switch seeking is less eccentric in some specific conditions? Could it be that for switch seeking in the weak sense all that is needed is an absence of any preference for the status-quo? In this paper we set out to explore the possibility that a considerable proportion of people are disposed to manifest switch seeking in a sort of decision states, characterized by being basically symmetric and simple to specify objectively.

The question is more than academic: If switch seeking of that kind was prevalent, it might turn out to be unhelpful or even counterproductive, not only for the decision maker but also for others, even for the whole community. A prominent example is the case of people, often called "noise traders", who rush to buy or sell stocks they hold based on insufficient evidence (e.g., Hughen & McDonald, 2005). It has been claimed that the activity of such noise traders accounts for much of the variance of stock exchange fluctuations. It seems reasonable to assume that noise traders are switch seekers that often act on the premise that they would be better off hitching a ride on another bandwagon.

The fact that their decisions prove to be often wrong indicates that they are guided, or even driven, by some fallacious reasoning of the kind known to be quite prevalent in human intuitive decision making (cf, e.g., Gilovich, 1991; Gilovich, Griffin & Kahneman, 2002; Kahneman , 2003; Kahneman & Tversky, 1984; Stanovich & West, 2000; Tversky & Kahneman, 1974, 1981) despite its deviation from normative models, notably EV (Bernoulli, 1954/1738) and SEU (von Neumann & Morgenstern, 1944; cf also Fishburn, 1983; Karmarkar, 1978; Machina, 1982; Markowitz, 1952; Savage, 1954).

The source of the fallacy in the reasoning that bears switching of that sort seems elusive. We explore the possibility that the culprit may be the fact that switching seems to be prescribed by a fallacious comparison of possible gain and possible loss, in a way that unfortunately might appear akin to what the normative procedure of maximizing expected gain. This is particularly well illustrated in the paradoxical *envelopes problem*.

1.1 The envelopes problem

Consider the following hypothetical problem: Suppose both you and another person are given sealed envelopes with amounts of money that neither of you knows anything about except that one of them is either twice as large or twice as small as the other one. Would you prefer to keep your own envelope or rather trade it with the other player?

The problem, in various variants (e.g., Brams & Kilgour, 1995; Christensen & Utts, 1992; Nalebuff, 1988; Samet, Samet, & Schmeidler, 2004; Sobel, 1994; Zabell, 1988; see review in Nickerson & Falk, 2006), is often termed the "exchange paradox". It seems paradoxical for the following reason: The player appears to have a rational basis to opt for trading envelopes, since by that she buys the chance (with probability 0.5) to obtain a double amount, while risking (with the same probability) losing only a half of whatever she has. However, since the *other* player has the same rationale for trading envelopes, the two are bound to find themselves after the exchange at an identical decision point, hence apparently motivated to keep trading envelopes over and over had they not remember what had led to the first exchange.

Since that behavior is apparently apt but really inapt, it seems paradoxical. Is it truly paradoxical? Nickerson & Falk (2006) argue that trading behavior in some variants of the problem is often justifiable given sound subjective assumptions, presumably tacit, about the distribution of possible amounts involved. Butler & Nickerson (2008) demonstrate that participants do trade at considerable rates when provided with explicit reasons to do so. Thus, even though people cannot explain very well the consequences of their behavior, it seems to be nonetheless adaptive in most naturalistic circumstances.

However, the particularly interesting part of the puzzle is not that some people, even ones coached in rational decision making, may be led or tempted to trade envelopes, sometimes for a sensible reason. It is rather that some normative rule people are advised to apply - namely, "decide between options by expected value of gains" - prescribes such a behavior even when it is patently absurd. Realizing its odd consequences calls, of course, for reconstruing the problem in a way that recovers symmetry and stability, which various authors offered.[1] Still, it is disturbing that even experts may be tempted to apply that normative rule in a misguided way. One may wish, then, to re-examine the premise that the applicability of that rule is universal.

Below we argue that the paradox results from applying that rule (namely, maximize expected gain) in a context that rather calls for another rule. We then discuss possible psychological accounts for such an error. Following that, we demonstrate a considerable propensity to switch in an experiment mimicking driving in a two-lane road.

1.2 The fallacy

It might be helpful to discuss the fallacy from the point of view of whoever is given the envelopes problem. We thus phrase the arguments below in the first person.

If the problem is regarded as "I have Y dollars, the other player has either 2Y dollars or Y/2 dollars", then the dilemma whether or not to agree to an exchange is at first blush answered in the positive, because the expected value of an exchange is greater than Y dollars ($0.5 \cdot 2Y + 0.5 \cdot Y/2 = 1.25Y$). This reasoning is based on the premise that though the assets in my envelope are not specified, they must amount to a definite sum of Y dollars, so that my uncertainty is just with respect to the state of envelope assets of the other player: Does he have 2Y dollars or Y/2 dollars? Framed somewhat differently, my puzzlement under that premise is which of the following asset compositions is the real one - (1) I have Y dollars and he has 2Y dollars, (2) I have Y dollars and he has Y/2 dollars. The consequence is the same: It appears advantagous to exchange envelopes.

However, actually my uncertainty is much greater. The symbol "Y" stands for my assets in the envelope. Normally that corresponds to a specific amount that I know. In this case, however, I do not know what that amount actually is, hence I do not have a tangible egocentric reference point. I should have noted that. Yet, referring to the amount in my envelope as Y, whatever the assets state of the other player is, lures me to assuming that I do have some definite and fixed, albeit unspecifiable, amount.

Since I do *not* know the amount in my envelope, a more appropriate way to state the problem, accordingly, is from an *acentric* point of view. The acentric formulation of the problem is the following: Postulating that the total amount of money is fixed (since otherwise the paradox does not arise; see McGrew, Shier & Silverstein, 1997), there are two possible states of the world – (1) player A (who happens to be me) has X dollars and player B has 2X dollars, (2) player A has 2X dollars and player B has x dollars. These states are equally likely.

In this formulation the uncertainty of any player with regard to her assets is the same as her uncertainty with regard to the other player's assets: She does not really have a definite amount, Y. Rather, she *either* has X or 2X, and so does the other player. That subtle yet crucial change in framing dispels the paradox, since now if I agree to exchange, I will gain X dollars if I am in state 1, and lose X dollars if I am in state 2. The expected value of my gain from an exchange is, thus, 0, which invites indifference between the options. Given typical conservatism, players would hence tend to select the passive option – do not trade envelopes.

So, the customary formulation arouses some fallacy of reasoning[2] that might elicit a futile behavior. Where does that fallacy come from? It must be due to some error in subjective framing, as many paradoxes are, but what kind of error?

1.3 Source of the fallacy

The fallacy described above is probably due in part to a disposition to view the world from an egocentric point of view, at least in the following epistemic sense: I know what I have; it is the state of the world outside me, or the intention of the other player, that I have uncertainty about.

We tend to forget that our own state might also be part of the segment of reality which we are uncertain about. Hence, we read Y as a symbol standing for a definite amount, whose actual magnitude is immaterial for making our choice. The reason that amount is denoted by a symbol must be taken by us to be aimed just for seeking generality. That reasoning presumably causes Player A to gloss over the crucial point: That although the quantity denoted by Y is definite, the fact that it may be either twice as large or twice as small as what the other player has renders it indefinite from *Player A's* point of view, which is utterly pertinent to her choice.

That neglect leads Player A to misrepresent the problem as being much like an ordinary bet, namely a choice between the status-quo (remaining with a sure amount, Y, namely with a zero gain) on the one hand, and a mixture of two changes in assets that are unequal in both sign and magnitude, Δ_1=2Y-Y= Y or Δ_2=Y/2-Y= -Y/2, on the other hand. Consequently, the decision maker presumes the gain/loss ratio is known to be greater than 1, as it is in many bet situations.

The crux of the fallacy is that it fails to take into account that in fact the status-quo itself is a mixture of two states of assets, X or 2X, and so is the status-quo of the other player. It is so, because for all the players know, the story is most plausibly the following: A fixed amount, Z (which equals 3X), has been divided in two - Z/3 (namely, X) and 2Z/3 (namely, 2X). Each of the players has even probabilities to have in her envelope either X or 2X. The magnitude of the difference between the states is thus *known*: |X|, hence actually the problem is a choice between the status-quo and a mixture of two changes that differ only in sign, Δ_1=2X-X= X or Δ_2=X-2X= -X. The symmetry follows.[3]

Another way of putting it is by comparing the expected value of one's own amount (0.5·Z/3 +0.5·2Z/3 = Z/2) with the expected value of the other player's amount, which of course is exactly the same. Again, the symmetry follows.

Note that the sample space is defined over the possible divisions of the overall amount, Z, *not* over the possible values of Z (for which the players seldom have any idea of, or very diffuse prior probabilities for at the best). Yet, the egocentric perspective regarding the exchange option as a mixture of two equally likely consequences, 2Y and Y/2 (where Y denotes the amount in one's own envelope), entails the latter sample space, since the above two consequences correspond to two possible values of Z: 3Y or 1.5Y.

The egocentric disposition is presumably enhanced when the player is permitted to know the exact amount in her envelope, namely that Y=$c where c is a known constant. However, the nature of the problem does not change just by knowing what Y amounts to. Knowing that does not add to the player any pertinent information, neither on the magnitude of Z nor on whether Y equals Z/3 or 2Z/3. What she does know is that both she and the other player have the same expected value for the amount in each one's own envelope ($.5 \cdot Z/3 + .5 \cdot 2Z/3 = Z/2$). As that value differs equally (by $|Z/3|$) from each of the amounts in the two envelopes, the expected gain (namely, the difference between whatever amount one actually finds in her own envelope and the one expected to be in the other player's envelope) is zero[4]. Hence, there is no reason to opt for exchange.

More generally, whenever two assets having unequal values, V_1 and V_2, are assigned at random to two persons, neither has any good reason to exchange assets, regardless of the magnitude of the ratio V_1/V_2. Rather, for what both players know, each is equally likely to gain $|V_1-V_2|$ or to lose $|V_1-V_2|$. Hence, the egocentric disposition must be fallacious, no matter what Z is and how it is divided, simply because it fails to admit the basic symmetry of the scenario.

Presumably, the fallacy is manifested more often under certain framings or conditions. Still, the fact that it is observed at all indicates the existence of the egocentric disposition. Without such a disposition, respondents would not normally fail to notice the complete symmetry of the situation, which is quite glaring once one gets to realize it.

The fallacy thus lies in people’s habitual perspective that, alas, proves to be wrong in the particular case. That is a common trap in paradoxical problems. Consider, for example, one well-known instance of Zeno's paradoxes, Achilles and the Tortoise: Though Achilles is ten times faster than the tortoise, if the tortoise had an edge (say, of 10m), Achilles would never overtake him, since by each time Achilles arrives at the tortoise's previous position, the tortoise would have advanced 1m further.

The lacuna in this logic is hard to spot, because the framing induces a perspective that makes us view this as a case of infinite regress – leading us to gloss over the simple truth that though a finite distance may be viewed as a sum of an infinite number of subdistances, it is still finite, hence traversable in finite time.

1.4 Extensions

Since that *egocentric disposition* is probably not restricted to this particular paradox, fallacies of this sort may be more prevalent than typically believed. It seems apt to denote them *the extended exchange paradox*. The phenomenon of noise trading may be a quintessential case. Another one may be the familiar phenomenon of queue switching (in the supermarket, on the road etc). Still another is the phenomenon of changing habits perceived as hazardous or beneficial to personal health.

Note that the fallacy in and of itself does not necessarily lead to overswitching behavior, since it may be counteracted by other predilections (e.g., conservativeness, the status quo bias, the endowment effect, loss aversion, concern about likely regret or disappointment). However, it clearly works to tip the scale toward switching.

Let us examine the three illustrative instances mentioned above, speculative as they are for the time being.

Life expectancy case. Consider, for example, the effect of a certain treatment (or a certain food, or any other substance that one can electively take) T on residual life expectancy E. Suppose that present evidence is equivocal. Some findings suggest that T prolongs E by the factor[5] R (R>1), while others suggest that it shortens E by the factor R. The balance of evidence leaves the two possibilities equally likely. Viewing the problem in this way does not result in any recommended conduct with respect to T.

Consider a person A who *does not* at the present take T. She might reason like this: Presently, without T, I have some residual life expectancy E (though I do not know its exact value). If I *opt for* taking T, I might (with probability 0.5) prolong it to R·E, or (with the same probability) shorten it to E/R. The expected E is, thus, 0.5· (R-1/R) ·E, which is clearly greater than E.

Now consider a person B who presently *takes* T. She might reason similarly: Presently, with T, I have some residual life expectancy E (though I do not know its exact value). If I opt for *quitting* T, I might (with probability 0.5) prolong it to R·E, or (with the same probability) shorten it to E/R. The expected E is, thus, 0.5· (R-1/R) ·E, which is greater than E.

Clearly, if both A and B reason the same, hence get to exchange treatments, their reasoning represents a logical fallacy, even though one of them may turn out to make the right choice. Of course, not all people are switch seekers. Some may be even switch aversive. But it is doubtful that they do so for the right reason.

Queue switching case. Now consider the phenomenon of queue switching. Simple common sense suggests that one should not be tempted by an advance in another queue (unless it is an exceptionally significant one), because such events are expected to occur by chance (and sometimes be due to errors of perceptual judgment; see Redelmeier & Tibshirani, 1999, 2000). Most people must be aware that they are just gambling by opting for a switch (except when it is clear that they can immediately materialize its fruit, like when using a temporary free nearby space in the alternate lane of the road just to overtake a car or two previously driving ahead of them in their own lane). The reason they often do that nonetheless may be due to an assumption made about that gamble - one that also might follow from applying the fallacy arising from the egocentric disposition.

The argument, explicit or tacit, presumably goes this way: If I am lucky, the rate of progress on the other queue will be faster than the rate on my present queue by a factor of R (say, 4 queuers per minute rather than 2 per minute). If I am rather unlucky, the former will be slower than the latter by a factor of R (say, 1 per minute rather than 2 per minute). The gamble seems to have a positive value.[6] Alas, it seems so from the perspective of any queue. It does not - to whoever does not wait in any of the queues.

Noise trading. Another instance of switch seeking seems to be the phenomenon of "noise trading", the activity of people who use to trade stocks they hold, typically in a bullish manner, based on insufficient evidence.

As noted above, noise traders must act on the premise that they would be better off switching to another stock, but the fact that they are often wrong indicates that they are driven by fallacious reasoning. A plausible candidate is the egocentric disposition.

A moderately rational noise trader may reason, much like the typical queue switcher: Stock A that I presently have is probably going to yield a return Y. How about the alternative stock, B, recommended by my broker? Of course, I may either win or lose by trading A for B. Yet, if I am lucky, the return yielded by B will be R times greater than Y, and if I am rather unlucky, it will be R times smaller than Y. Since the deal seems to have a positive value, let me go for it.

Actually, a similar line of reasoning could be made by present holders of B. But that, unfortunately, is glossed over by our noise trader.

From experience, the noise trader may later learn the hard way that she has been wrong. If she does not attribute that to bad luck, she might grow to lose faith in her reasoning, then become more aversive to further engaging in such decisions (Ido Erev, personal communication). It is a long shot, though, that she would put the blame on decision rules she has been taught. However, that would probably have been the soundest conclusion.

A moderating factor. In all these instances, unlike in the envelopes problem, switch seeking must considerably depend on the extent to which it may be reasonably assumed that the dominant option in the present continues to dominate the alternate one in the immediate future. Yet, a considerable disposition to switch perhaps exists even when, just as in the envelopes problem, the person does not even know which is the dominant option at the present.[7]

2. The present study

Empirically testing the presence of the extended exchange paradox in real-world situations calls for studying the prevalence of switching behavior. That in turn requires a paradigm in which decisions, akin to those being taken in some real-world situation, may be described as switching from, or staying at, some status-quo defined on a trial-by-trial basis. For that purpose, we studied how frequently participants tend to switch lanes in an experimental setting simulating the state of a driver on a two-lane, unidirectional road.

An experimental support for the existence of a propensity to switch in the face of slim evidence for some merit from switching requires demonstrating dependence on variables that appear to be related to such evidence, to rule out mundane accounts (e.g., that participamts do it out of curiosity, boredom, or a bias in favor of probability matching in case of indifference).

We conducted an experiment to seek such support. Participants were not asked to actually drive. They were rather told that they would be able to watch their journey back home from work on a monitor screen as it proceeded. At any given trial, the participant watched her car traveling along a 1-km segment of the way in one of the two lanes. Cars in each lane were moving at about the same speed, where the speed in one of the lanes was twice as fast as the speed in the other lane. Though speed assignement on each trial was totally random, the participant was not informed about that. At the end of each trial the participant selected the lane for the following trial by choosing whether or not to switch the lane her car was driving in.

Participants were encouraged to try minimizing travel time. A straightforward strategy that they could use in those circumstances, had they supposed that speed assignment was not perfectly random, may be termed *Apparently-Rational Switch Strategy* (*ARSS* for short): Switch following a "slower" trial; do not switch following a "faster" trial. We tested the hypothesis that participants tend to use that strategy, implicitly or explicitly, by studying how frequently lane switches were made, how switching behavior depended on traffic variables, how participants perceived switches, and whether the tendency to switch lanes related to held positions about the possible merit of switching.

2.1 Method

2.1.1 Participants

Ninety-eight undergraduate psychology students at the University of Haifa participated for course credit. A monetary performance-dependent prize to a single winner of a lottery was promised as well. All participants had normal or corrected-to-normal vision.

2.1.2 Apparatus and setting

Stimulus presentation and data acquisition were controlled by an O2 SiliconGraphics computer. Stimuli were presented on the screen of a 22-inch computer display. The participant sat in front of the monitor at a viewing distance of about 60 cm. She responded by pressing on a designated keyboard key (the "Z" or the "/" keys) with the index finger of her dominant hand.

2.1.3 Stimuli, design and procedure

Each participant was run individually in one experimental session that lasted about an hour. She was asked to imagine that she was driving from work to her home during heavy traffic in a one-way two-lane highway, and was told that she would be able to watch her journey on the monitor screen as it proceeded, as viewed from a camera that moved synchronically with her car. The display showed car icons (from a top view) "moving" on a limited stretch of that road. At any given moment, cars in each lane were moving at about the same speed, where the speed in one of the lanes was twice as fast as the speed in the other lane. The side of the faster lane (right, left) was randomly varied between trials.

One of the cars, distinct from the other cars by color and size, was designated as the participant's car. It was suggested to her that she is expected to try to arrive home as early as possible (by mentioning that "since the route is fixed, the faster you travel, the earlier the experiment ends"), and could be rewarded (a monetary prize proportional to her speed of "reaching home", namely to the inverse of her travel time, in case she would win in a lottery conducted after all participants were run). She had no direct control on the speed of her car, but could try to affect it indirectly by choosing the lane on which it would "move" in the following trial. She was also told that on each trial she would be watching the car traveling along a 1-km segment of the way, and that at its end she would have to decide whether or not to switch lanes during the following trial showing the subsequent segment.

The cars displayed were heading toward the top of the screen. Yet, the participant's car was never seen moving up; it was rather stationary at the bottom of the screen, while the surround (marked by icons of electricity poles on the road margins) "moving" toward the bottom. All other cars within the same lane were also seen as basically stationary (except for slight jitters due to introduction of very small, random perturbations in either of the two directions - top and bottom). Cars in the other lane, in case they were visible (see below), were seen "moving" up when their speed was higher, or down when their speed was lower.

The lanes appeared in the middle of a white rectangle (21 cm in height and 10 cm in width). Each lane's width was 2.5 cm. Cars "moved" at the middle of their lane. The cars measured 1.5 cm in height and 0.8 in width. They appeared in various colors (yellow, purple, gray, red, blue, green and azure). The participant's car was somewhat larger (1.8 cm in height and 0.8 cm in width) and black. It was invariably located at the bottom of the lane (with its back 4 mm above the bottom of the rectangular frame). The lanes were divided from each other by a black solid line. A red solid line was used to mark the margins of the lanes. The number of cars visible in each lane at any given trial was either 3 or 4. The icons of the electricity poles at the margins were drawn as circles (2 mm in diameter). The spacing between poles was 10.5 cm, so that at any given moment one or two poles were visible on each road margin. An illustrative display is presented in Figure 1.

Insert Figure 1 about here

As noted above, in each trial the speed of cars in one of the lanes was twice as fast as in the other lane, where the side of the faster traffic was varied randomly. The sequence of actual speeds in the slower lane was determined by randomly sampling from a tripartite uniform distribution (with an overall range of 5-90 kmh, with probability of 0.1 for the subrange 5-15 kmh, probability of 0.7 for the subrange 16-45, probability of 0.2 for the subrange 46-90, yielding a mean of 30 kmh). The speed in the faster lane was twice as high (yielding a mean of 60 kmh). The length of a trial varied between 7.3 sec to 150 sec.

On each trial except for the first one, the side of the participant's lane (right vs left) was determined by her response (switch or stay) following the immediately preceding trial. The participants were also given online feedback about the percentage of the total distance to home already travelled by the time, by means of a incrementally growing thin green stripe at the bottom of the screen. The full length of the stripe (at the end of the road) was 9.5 cm.

At the end of each trial, except for the last one, a notice was presented at the bottom of the screen, asking the participant whether she opted to switch lanes or to stay at the same lane on the subsequent trial. The participant were to react twice: First, make a quick decision (within 1 sec) indicating her gut, tentative intent; second, within 3 more sec, give her final, more deliberate decision. Responses were given by pressing on one of two designated keyboard keys.

The number of trials in the experimental session was 70. The first ten ones were practice trials. In the first five practice trials, the participant was asked to just watch the traffic in the two lanes without responding. In the following five ones, participants were asked to also respond. The following 60 trials were considered as experimental, hence were analyzed.

The 47th experimental trial was special, in that no cars (including the participant's car) were shown on the screen, and an announcement was displayed saying that the camera was temporarily out of order, but that the participant was supposed to respond despite that.

The order of relative speeds over trials was manipulated between participants. Since the random assignment of relative speed to trials we initially used (on the first group of 63 participants), coupled with the response patterns of many participants, turned out to generate reinforcement of the ARSS strategy in more than half of the trials of the first half of the sequence (about 58.5%), but less than half of the trials (32%) - in the second half, we used another group (of 35 participants) for which the order of relative speeds over trials was reversed. Henceforth, that between-participant order variable (ARSS-encouraging vs ARSS-discouraging, respectively) is being referred to as trial history.

At the end of the session, following the last trial, participants were asked to fill out a pen-and-paper questionnaire about how they perceived the task and their behavior while performing it (see Appendix A).

2.2 Results

The data of four participants were excluded for having a response percentage smaller than 2/3. The data of all other ninety-four participants were analayzed. Mean response percentages and mean switch rates were calculated separately for immediate responses and for final responses, in both ordinary trials and the single no-feedback trial.

Switch rates were considerably far from negligible. The 99 percent confidence limits for mean switch rate in final responses are 0.401- 0.467. The 99 percent confidence limits for mean switch rate in immediate responses are 0.415 – 0.492. That was pretty much the case in the no-feedback trial as well: The Agresti-Coull binomial proportion confidence limits for the proportion of participants who opted for switching in final responses are 0.238-0.511. The per-participant correlation between immediate and final responses was generally high, with a median of 1.0 (mean=0.916). All the following analyses are on final responses in ordinary trials.

To examine whether participant behavior relates primarily to relative speed of traffic in her lane (twice or half of the speed in the other lane) or to the exact value of speed just in her lane, Pearson correlation coefficients were calculated for each participant between (a) her response (switch=1, stay=0) and the speed of traffic in her lane on that trial, as well as between (b) her response and the sign of the ratio of that speed and the speed in the other lane (twice as fast, twice as slow).

The overall percentage of participants (out of just those 65 ones who had both correlations negative) for whom the correlation between relative lane speed and response was higher than the correlation between actual speed and response was 89.2%. Percentages were not significantly affected by trial history, $p=.39$ by Fisher exact test. Thus, participants tended to switch predominantly following trials in which the traffic speed in the participant's lane was slower.

Mean correlations, as a function of trial history and correlation type, are presented in Table 1. Not surprisingly, in view of the results presented above, the mean correlation between response and actual speed was found significantly lower than the correlation between response and relative speed, $F(1,92)=4.44$, $p<0.05$, $MSE=0.263$. Yet, here the main effect of trial history was also significant, $F(1,92)=39.74$, $p<0.0001$, $MSE=0.022$, as was the interaction term, $F(1,92)=8.61$, $p<.005$, $MSE=0.022$.

Insert Table 1 about here

To obtain a more elaborate, fine-grained account of participants' switching behavior, we looked for further contingencies. Participant behavior could, of course, be arbitrary. Yet, several types of systematic patterns, or at least dispositions, might be anticipated.

One conceivable pattern of participant behavior is state-indifferent perseverance, in turn branching in two possible subpatterns – one in which the participant almost always switches, the other in which she almost never switches, in both cases regardless of what the present state seems to be and of what her previous choice has been. If not due simply to low collaborativeness of participants with experimental requirements, the former (namely, almost always switch) might arise from compulsive switch seeking, and the latter (namely, almost never switch) – from apathetic/ fatalistic disposition. On the other hand, both could be moderately reasonable strategical choices, in case the participant somehow had reason to believe that there was no way to optimize behavior in view of the high uncertainty about the speed ratio. Indeed, given the random allocation of speeds in the design, the expected number of successes (trials in which the speed in the participant's lane is faster following her response) is the same in both strategies – very close to 50%.

It is more interesting to examine patterns that indicate some contingency of response on present state or very recent history. Though no strategy could be better than another in this study, because speed allocation was administered totally at random, participants might have considered some strategies to be optimal as long as they did not realize the futility of that in view of the randomness (which may be hard, since people tend to perceive illusory correlations; see Chapman, 1967).

We hypothesized that many participants would apply frequently, albeit perhaps not systematically, the strategy termed above ARSS (amounting to: Switch following a "slower-trial"; do not switch following a "faster-trial").

Though it is not absolutely clear that participants explicitly, much less deliberately, used the ARSS strategy, the behavior of most of them suggests that they applied it, explicitly or implicitly, very often. To examine whether their behavior in that respect changed throughout the experiment, for the sake of the analysis we broke down the session in two subsessions – the first half, including trials 2-29 (namely, excluding the first trial, behavior in which could not be related to any preceding events, which we do in some of the analyses reported below), and the second half, including trials 30-60 and excluding the no-feedback trial and the subsequent one (numbers 47 and 48).

The initial analysis was conducted using a three-way ANOVA on mean switch rates, calculated for each participant in each cell of a 3X2 factorial design, with trial history (ARSS-encouraging, ARSS-discouraging), participant lane speed (faster, slower) and sub-session (first half, second half) as factors. The results are presented in Table 2A. The main effects of sub-session, as well as the history were not found significant, $F(1,92)<0$, and $F(1,92)=3.00$, $p<0.09$, $MSE=0.054$, respectively. On the other hand, the main effect of participant lane speed was found significant: Mean per-participant switch rate was found much larger in "slower" trials (0.620) than the corresponding mean in "faster" ones (0.226), $F(1,92)=52.68$, $p<0.0001$, $MSE=0.203$. The former was, furthermore, found to significantly exceed chance level, Binomial p=0.0105.

The dual interaction between participant lane speed and trial history was found significant, $F(1,92)=6.47$, $p<0.01$, $MSE=0.203$. The effect of participant lane speed seems to be more salient in the ARSS encouraging condition. No other interaction was found significant or close to significance.

Insert Table 2 about here

Inspecting the list of observed per-participant rates in all experiments, four types of participants were discerned: (a) *ARSS devout followers* – those who manifested ARSS-compatible behavior in each trial, (b) *ARSS preferrer*s – those who manifested ARSS-compatible behavior in at least 90% of the trials, but not in each, (c) *Stay preferrers* – those who opted for staying in at least 90% of the trials, (d) *Others*. The numbers of participants in this experiment falling into each of those types were 16, 13, 1 and 64.

Thus, almost a third of the participants were predominantly or entirely consistent in following ARSS. Almost all the rest (categorized above as type d) were less consistent, but on the whole they fairly often switched, and in addition manifested behavior that was compatible with ARSS more frequently than ARSS-incompatible behavior. As can be seen in Table 2B, across those other participants, the mean per-participant rate was higher in "slower" trials than in "faster" trials. The ANOVA conducted only on the data of those participants yielded a significant main effect of participant lane speed $F(1,62)=10.94$, $p<0.001$, $MSE=0.094$. No other main effect or interaction was found significant or even close to significance

Switch rates and their complements may be reframed as rates of ARSS-compatible responses. Table 3A presents those rates for all participants; Table 3B presents them for just the 64 "other" participants. In an ANOVA conducted on the rates in the former table, significant main effect of participant lane speed was found, $F(1,92)=41.33$, $p=0.0001$, $MSE=0.054$. The main effect of history was also found significant, $F(1,92)=6.47$, $p=0.01$, $MSE=0.203$. The interaction between history and participant lane speed was close to significance, $F(1,92)=3.00$, $p=0.09$, $MSE=0.054$. No other main effect or interaction was found significant or close to significance. In the ANOVA conducted on the rates in the latter table, significant main effect of participant lane speed was found, $F(1,62)=52.41$, $p=0.0001$, $MSE=0.058$. No other main effect or interaction was found significant or close to significance.

Thus, it appears that participants tend to stay in the faster lane more than they tend to switch from the slower one.

Insert Table 3 about here

To examine whether there were sequential dependencies between consecutive trials in data obtained from participants whose behavior did not seem to be determined by a single or predominant strategy, namely the 64 "other" participants, we conducted two further three-way ANOVAs on mean switch rates of those participants.

In the first one, the three factors were history, participant lane speed in trial i (namely, the one at the end of which the participant made the response) and participant's behavior in trial i-1(switch, stay). Whereas participant lane speed in trial i was, consistent with the above analyses, found to have such an effect, $F(1,62)=11.00$, $p<0.002$, $MSE=0.162$, and the effect of behavior in trial i-1 was also significant, $F(1,62)=10.68$, $p=0.002$, $MSE=0.088$, the pairwise interaction was not found significant ($F<1$), which indicates that there was no discernible sequential dependency: Participants' behavior was *unaffected* by success in the recent trial. The main effect of history as well its interactions with the other variables were not found significant ($F<1$).

In the second ANOVA, the three factors were history, participant lane speed in trial i and ARSS-compatibility of participant's behavior in trial i-1 (compatible, incompatible). Whereas participant lane speed in trial i was found, as in the preceding ANOVA, to have such an effect, neither the main effect of ARSS-compatibility in trial i-1 was found significant, $F(1,62)=2.00$, $p=0.17$, $MSE=0.060$, nor was the pairwise interaction, $F(1,62)=1.87$, $p=0.18$, $MSE=0.100$. The main effect of history as well its interactions with the other variables were not found significant or close to significance. Again, there was no discernible sequential dependency.

On the other hand, it could well be that there were higher-level dependencies: Though behavior did not seem to change much over the trial sequence, participants might have been instrumentally conditioned in early trials to later favor or retract some response strategy. To further examine whether or not participants were in any way influenced by experience during previous trials, mean per-participant stay rates in "faster" trials, mean per-participant switch rates in "slower" trials, as well as mean per-participant rates of ARSS-compatible responses, across all trials in the second half, were correlated with per-participant mean differential reinforcement rates of the two possible modes of participant's behavior (ARSS-compatible, ARSS-incompatible) in the two types of participant lane speed (faster, slower) across all trials in the first half.

The results presented in Table 4A indicate a salient effect of history. In the ARSS-encouraging condition, we found a fair amount of dependence between the participant behavior and her previous-trials experience. However, across participants most such individual influence effects were cancelled out, resulting in no significant effect of sub-session. ARSS-compatibility did not increase in the second half (and actually appeared to have somewhat decreased; see Table 3 above). That could not have been due to paucity of trials in which ARSS-compatible behavior was reinforced (by landing on the faster lane), because the percentage of such trials was about 50%. It is more likely to rather reflect some frustration of participants who expected, or at least hoped, that following ARSS would gain them a higher rate of success.

Insert Table 4 about here

The pattern in the ARSS-discouraging condition was different. The results indicate no significant dependence between the participant behavior and her previous-trials experience. The reason, as it seems on closely examining participants' behavior, is probably that scarce reinforcement of ARSS (on the average, 27% of the relevant trials in the first half of the session) was *not* sufficient for deterring quite a few participants from continuing to frequently follow it.

To illustrate, one participant followed ARSS in 100% of the trials at the second half of the session despite having been reinforced only in 30% (8 out of 27) of his ARSS-compatible trials in the first half. Another one followed ARSS in 67% of the trials at the second half despite having been reinforced only in 9% (1 out of 11) of his ARSS-compatible trials, but as much as 50% (7 out of 14) of his ARSS-*in*compatible trials in the first half. That extreme disregard of consequences must have been mediated by some wrong intuitive reasoning, since it is incompatible with theories of implicit conditioning (see, e.g., Kileen, 1994). Such participants may have been attributing their lack of success to bad luck. Ironically, though they were right about that, unfortunately that served as a rationalization helping them to hang on to overtrusting the illusion that the ARSS strategy must be a reasonable bet.

Since no significant learning in the ARSS-discouraging condition was observed, the six correlations were compared to the same correlations in the ARSS-encouraging condition. Four of them (except for the two between percent trial in "faster" and "slower" trials in trials 30-60 and between differential reinforcement rate in trials 1-29 in which speed in participant's lane is faster) were found significantly larger (with *p*<.05 at least in a two-tailed Z test) in the ARSS-encouraging condition.

This difference could be due to the small rate of consistent following of ARSS in the ARSS-discouraging condition at the start, that probably acted to avert considerable negative feedback, which in contrast many participants in the ARSS-encouraging condition did receive. Such a paucity of consistent data about contingency may result in opting for random choice (see Erev & Barron, 2005). Alternatively, it may be more plausibly due to behavior that persists irrespective of apparent success. It seems that positive feedback for ARSS stregthened it considerably more that negative feedback weakened it. This suggests that participants may have misperceived the situation, entertained some overoptimistic idea about the best way to deal with it, or perhaps both.

In any case, when the correlations were calculated across those two conditions (see Table 4B) that differed just on the order of assignment of relative speed to trials, all six correlations were found to be highly significant.

To compare within-session trends in following ARSS in both history conditions, we calculated cumulative ARSS-compatible rates, by trial 12, trial 30 and trial 60. The cumulative rates in the ARSS-discouraging condition were 0.660, 0.605 and 0.608, respectively. It is not at all clear that the smaller ARSS-compatible rate there is due in any way to the extent of learning from previous trials experience, since the rate decreased no less in the ARSS-encouraging condition 4 - 0.764, 0.744 and 0.728, respectively.

In sum, the main features of the results - considerable mean switch rate, high mean ARSS-compatible rate, as well as the dependence of both on participant lane speed – were demonstrated in both groups. This suggests that switching behavior is fairly robust and not highly sensitive to reinforcement history.

RT analyses

Two one-way analyses, GLM (due to the presence of missing cells) for immediate responses and ANOVA for final responses, were conducted to compare mean RTs for the two possible responses (switch vs stay), in each of the two immediacy levels.

Interestingly, mean RT of switch responses was found significantly smaller than mean RT of stay responses in both levels. In immediate responses, the means were 619 ms and 651 ms, respectively; $F(1,92)=43.70$, $p<0.0001$, $MSE=1160$. In final responses the means were 2487 ms and 2520 ms, respectively; $F(1,28)=8.34$, $p<0.01$, $MSE=2016$. It thus seems, that switches was in a way the default option, presumably because participants were inherently disposed to choose it.

In passing, the practically-equal effect sizes, coupled with the very high correlation between immediate and final levels, suggest that the final response was in the great majority of cases just a delayed replica of the immediate response, and were not preceded by plenty of second thoughts.

Questionnaire data

Participants' replies about their own frequencies of switching (first two questions) indicate that they were fairly well aware of their behavior: The mean self-estimated number of switches of those participants who reported of having switched frequently was 26.82, significantly more than 13.48, the corresponding mean of participants who reported of *not* having switched frequently, $F(1,64)=18.36$, $p<0.0001$, $MSE=157.65$. The objective means of number of switches for those two participant types, 27.44 and 22.02 respectively, were also significantly different, $F(1,80)=12.41$, $p<0.001$, $MSE=47.58$.

A large majority of participants (77.4%) reported of being able to describe the characteristics of the state in which they had opted for switch. The verbal replies (of 73 participants whose phrasing was clear enough) about that, together with the arguments pertinent to switching, were sorted by us in four categories – (a) *absolute*: traffic in participant's lane seemed slow, (b) *relative*: traffic in participant's lane was slower than in the alternate lane, (c) *history*: the participant was applying lessons from previous trials, (d) *other*: a variety of reasons that do not fall in the former three categories. The corresponding frequencies were 5, 44, 5 and 19, respectively.

Mean number of switches was found to significantly differ among the participant types defined by those four categories, $F(3,69)=2.72$, $p=0.05$, $MSE=44.11$, where the only significant difference between specific types is that between "absolute" and "history". A significant effect of participant types was found also with respect to mean ARSS-compatible rates, $F(3,69)=5.51$, $p=0.002$, $MSE=0.049$. Here, "relative" significantly differ from "absolute". In addition, those categories were found to be associated with our typology of participants' strategies: Out of the 24 participants whose behavior seemed to follow ARSS at least in 90% of the trials, 92% gave a reply categorized as "relative"; out of the 49 participants whose behavior was typed as "other", only 45% gave that reply, $\chi^2= 14.72$, $p<.0001$. Accordingly, it seems plausible that at least half of the participants were led to frequently follow ARSS, because it appeared to them a promising way to land often on the faster lane.

Out of 90 participants who responded to the two questions about the envelopes problem, 77 replied twice that they would not exchange envelopes. Out of the 13 others, six replied twice that they would exchange envelopes, five replied so only to the post-argument question (apparently having been persuaded by the argument), and the remaining two replied so only to the pre-argument question.

Interestingly, out of the 13 participants who at least once replied that they would opt for exchanging envelopes, 54% (namely, seven) followed ARSS at least in 90% of the trials, a percentage that is much higher than the corresponding percentage among all other 77 participants – 26%, $\chi^2=4.11$, $p<.05$.

3. General Discussion

A major finding of this study is that many participants, far more than a small minority, switched lanes frequently, despite having no objectively sound reason for doing that. Another major finding is that they did that substantially more in trials in which they found themselves in the lane in which traffic was moving more slowly, which suggests that they were trying to do the best they could, in their view, to meet the goal implied in the instructions. A number of them appeared to follow the ARSS strategy, prescribing switching in such trials and only in them. Some other participants appeared to have also practiced that strategy, yet less systematically.

It goes without saying that no strategy was really optimal. Indeed, participants who predominantly or entirely followed ARSS actually achieved a rate of success (namely, landing on the faster lane following the response) that was practically the same as the mean success rate of other participants (47.6% vs 47.3%, respectively, $F<1$).

Explaining participants' behavior is far from straightforward. Despite the paucity of information participants could rely on at the start to figure out how to optimize their chance to “reach home” as soon as possible, they were explicitly told in the instructions, and anyway must have become pretty early aware of the basic symmetry between lanes, as well as of their own uncertainty whether a switch would get their car into a lane that was twice as fast or twice as slow as the lane they were in at the time. Furthermore, since lane speeds were determined randomly, the objective rate of positive feedback the participants received was not very high. Thus, they had no good reason to suppose that since switching was an offered option, it must have been a smart move to apply.

This is also true of participants who adopted a fixed strategy, notably ARSS. For those participants, their strategy dictated a rate of positive feedback very close to chance level across all trials (actually, it turned out slightly less). Hence, at least as they have acquired some experience, they should have been realizing that the distribution of lane speeds over trials was not determined in any nonrandom pattern (e.g., a systematically faster lane, a simple alternation rule) that they could count on, and that even if there was occasionally a temporarily faster one for a run of a few trials, their chance of landing on it again was meager. In view of that, despite the well-known human handicaps in correctly perceiving randomness (e.g., Kubovy & Gilden, 1990; Lopes & Oden, 1987; Nickerson, 2002; though strong correlations are detected relatively fast, see Kareev, 2000), they must have not developed high expectations about discovering a heuristic, much less a formula, for beating the system. Judging by the questionnaire data, nobody did.

The behavior of most participants was, nonetheless, sensitive mostly to a trial-specific stimulus - lane speed. In that respect, it resembled behavior of participants in some other paradigms involving more artificial tasks, such as gambling which of two boxes contained a reward (Avrahami & Kareev, 2011). On the other hand, unlike participants in the latter study (in which decisions were *not* referred to the status quo, and one alternative was preferrable in the long run), behavior of participants in the present study was *not* affected by the success of their choice in the most recent trial. No sign of interaction between participant lane speed in trial i and participant's behavior in trial i-1 was discerned. Rather, our participants typically based their choice simply on lane speed in trial i.

Thus, despite likely skepticism about the chance of discovering a useful heuristic, most participants must have been surmising (or implicitly been driven to act as if they were surmising) quite frequently that somehow they had more to gain than to lose from a switch following a "slower" trial. The big question is *why* they felt that way.

The case for adopting the ARSS strategy presumes some extent of short-term stationarity of lane speed beyond that expected from a basically random process (which is, of course, a likelihood of 0.5 for getting the same relative lane speed as before), reminiscent of the belief in "hot hand" runs, prevalently biasing perception of sequence randomness (see, e.g., Gilovich, Vallone & Tversky, 1985). Yet, whereas it seems *a-priori* reasonable to surmise that "hot hand" accounts for some variance in psychomotor behavior, as an hypothesis about disparate random sequences, whose generating process is unknown, it is not less arbitrary than the converse hypothesis - alternation bias. Since participants were not provided in the instructions any clue that the process was not random, but most of them (actually, all but one) tended to prefer ARSS from the start, they must have been drawn to that preference by making a wrong guess about the purpose and design of the experiment and/or by importing their existing model of road behavior on the highway.

True, participants' predispositions at the outset were somewhat changed later by apparent contingencies generated by perceived deviation of the subsequence at the start of the session from their imaged model of perfect randomness. Yet, the mean rate of ARSS-compatible responses exceeded chance level even when experience has not seemed to reinforce it. Ironically, lack of reinforcement appeared to have somewhat lowered adherence to ARSS in "faster" trials. That was presumably due to a tendency to opt for switching that was fairly strong regardless of whatever might have been learned from previous trials: Mean switch rate across all participants in the ARSS-discouraging group was 0.393, just slightly smaller than in it was in the ARSS-encouraging one – 0.438.

Furthermore, switching was found to be opted for even in the face of *full* uncertainty: The appeal to switch was considerable (mean switch rate of 0.390) also in a control experiment that we conducted (not reported here), in which participants were not shown the traffic in the other lane. Even in the no-feedback trial reported above, in which no speed information at all was available to the participant, mean switch rate was almost as high as it was in ordinary trials, 0.5 for participants who mostly or entirely followed ARSS, and 0.316 for participants in the "others" category. This seems fairly consistent with the percent of respondents prefiguring they would switch queues in the supermarket (see Footnote 7).

Thus, presuming short-term stationarity was not found to be a necessary condition for switching. Actually, it could neither have been a *sufficient* condition: Assuming that lane speeds on trial i+1 are somewhat more likely to remain as they were on trial i than to change does not, of course, eliminate the risk that they do change. The aversion to such risks very often gives rise to a bias favoring the status quo. It, thus, appears that there must have been some other reason for being tempted to switch.

That the option of switch had a pre-rational appeal receives some support from the finding that mean reaction time of switch responses was significantly lower than the corresponding mean in stay responses, in immediate and final responses alike. Conceivably, a decision to stay took some extra time, perhaps required to counteract the spontaneous impetus to switch. Unlike in real road driving, in which that impetus is often restrained by caution since the loss is most often considerable (and there is some small increase in risk of accident from switching into another lane), in the artificial environment of this experiment the loss must have appeared trivial, so that the impetus to switch could be more openly manifested. This seems to be a virtue of the design, since the issue is how strong that impetus is, not how frequently it actually triggers switching (and not just in highway driving, which is after all just one of the possible arenas in which switching is considered).

All that lends some support to the hypothesis, that people, at least some people, often seek switching in default of any subjective evidence for stationarity and sometimes even in the face of cumulating evidence for the contrary. It seems that anyhow the crucial consideration is not any illusory belief about beating uncertainty. To illustrate, even when a basketball coach tends to attribute a run of missed shots by a shooting guard to a "cold hand", he must know full well that there is no guarantee that the player sent in to replace the unluck one would fare better. If he nonetheless opts for replacing the "cold-handed" guard, he must probably believe that there is more to gain than to lose from doing that. That must be a-fortiori true in case there is no apparent evidence for stationarity. People may, thus, find switching attractive for some reason that transcends whatever evidence they might have, or its lack thereoff, about uncertainty.

What *is* that reason? It is all too easy to ascribe that preference to a long history of reinforcements. As argued above, it is not at all clear that switching behavior could in general be truly beneficial.

In any case, the answers to the questionairre items suggest that most of our participants were not acting instinctively, much less whimsically. Rather, they were presumably giving some thought to the problem, even those ones who did not end up locked on a single strategy. Granted, their solutions must have been far less than articulated. So is, however, most of human informal decision making. That does not daunt students of behavior from inferring the existence of implicit rationales (SEU, for one) underlying the behavior in question.

So, is a propensity to switch actually supported by the data?

One might contend that the existence of a *tendency* to seek switching has not been substantiated here, since the overall mean switch rate was found to be less than 0.5 (surmised to be chance level). Alas, there are two faults in that argument.

The first one is that it seems to ignore a very plausible possibility: Like any other tendency, switch seeking may not be general across people, neither non-circumstantial. In the Introduction we suggest that switch seeking is a trait rather than a universal phenomenon. It is quite possible, even likely, that while many people normally avoid switching, perhaps due to the status quo bias (possibly enhanced, in the case of road driving, by the small risk in lane switching), some others are habitual (even persistent) switch seekers, and yet others are occasional (presumably circumstantial) switch seekers. In this experiment, 31% were found to be quite consistent switch seekers. Hence, mean switch rate across *all* participants is anyhow not expected, much less predicted, to exceed chance level.

A more fatal fault in that argument is that it posits chance level switch rate to be 0.5 across *all trials*. Yet, the hypothesis being tested here is obviously *not* that switch seeking is committed out of compulsion irrespective of circumstances. We conjecture that even habitual switch seekers mostly do not crave switching quite irrationally, but are rather guided by some pseudo-rational argument, conscious or preconscious, leading them to frequently opt for switching. Furthermore, they do not do that always, only whenever they feel for some reason that their state could have been better off. In the reported experiment, for example, ARSS followers switched lanes only or mostly following a trial in which they found themselves in a lane in which traffic was relatively slow (which they probably anticipated to persist to the subsequent trial). The figure pertinent to our hypothesis is, thus, their mean rate of switching in *that* state. It was found to be 0.99 for ARSS followers across both kinds (devout followers and preferrers). Furthermore, even participants having no discerned behavior pattern opted to switch significantly more often after being in the slower lane. Across all participants, the mean rate of switching in that condition was 0.62, significantly higher than 0.5.

Thus, it seems plausible that many people have a considerable tendency towards switch seeking, and others are tempted to switch occasionally when that strategy appears to them apt.

Were participants manifesting the egocentric disposition?

In the present case, the implicit rationale for surmising that there is more to gain than to lose from switching could be the false notion that the gain is literally greater, due to a pseudo-rational argument of the sort shown above to follow from the extended exchange paradox, emanating in turn from the egocentric disposition. A participant need not be acquainted with SEU or any other formal decision model to have tacitly reasoned in about the following manner: If I am lucky, the speed in the other lane is going to be in the next trial faster than the speed in my present lane by a factor of 2. If I am rather unlucky, the former will be slower than the latter by a factor of 2. So, supposing the forthcoming speed in my present lane is S_0, the speed in the other lane is going to be either $2S_0$ or $.5S_0$. Since $|2S_0-S_0| > |.5S_0-S_0|$, this must be a risk worth taking.

As we pointed out in the Introduction section, people may behave as if they are implicitly guided by such a rationale without having explicitly considered it as such, much less having deliberately used it as an apparently reasonable basis for opting to switch, let alone having plainly calculated the expected gain. The mistaken intuition that there is twice as much to gain from switching than there is to lose may suffice. Yet, the supposition that "twice as fast" buys one more than what she could lose in "twice as slow" is not only wrong mathematically, predicating that the objective is to minimize travel time rather than maximize one’s rate of progress (see Footnote 6). Its main fault is that it is *logically* fallacious.

As argued above, the fallacy results from thinking solely in terms of gains and losses. Gains and losses are defined with respect to a given status quo or to some other reference point. Actually, however, when the participant is to make her choice of lane for the subsequent trial, i+1, she does *not* have any subsequent-trial-pertinent reference point. Though she might somehow take into account, or be otherwise affected by, her speed in trial i (to the extent that she assumes short-lived stationarity), that speed is irrelevant, because she cannot have any idea how faster or how slower could her speed in trial i+1 be. All she does know about the subsequent trial is that like in any other trial, the speed in one of the lanes is going to be twice as large as the speed in the other lane. If she happened to end up in the former, she would obviously move twice as fast as she would have had she ended up in the latter. If she rather happened to end up in the latter, she would move twice as slow as she would have had she ended up in the former. Those are not two separate reference-point-oriented *outcomes*, as gain and loss must be, but rather two possible-outcome-oriented *framings*.

The finding that switching was more frequent in "slower" trials suggests that participants presumed some level of stationarity. Yet, the finding that it was exhibited quite often in "faster" trials as well (especially by "other" participants), or in no-feedback trials, supports the claim that they were guided by the argument following from the extended exchange paradox.

The error in hoping to gain from lane switching should be evident, perhaps even glaring, to any reader at this point. But is it as evident to a person naïve about the problem, skilled as she may be in normative decision making?

Interestingly, the same fallacy would be manifested in the envelopes problem, if the customary maximizing-expected-gain rule were applied to it from the viewpoint of egocentric framing. It has not been demonstrated that people are actually likely to commit the fallacy-originated decision when presented with the envelopes problem in its classic form (4 out of 30 participants in Butler & Nickerson, 2008, Exp 1, Game 6; 8 out of 90 - in the questionnaire data we collected). They must be counting on their intuition in that case more than on maximizing-expected-gain reasoning, even when they are moderately versed in applying it. Perhaps it helps that in that form of the problem there is no seeming clue that the person holding the other envelope may be better off.

The case may be different in situations involving choice between queues. Gambling on switching queues in a supermarket, a more mundane dilemma, perhaps seems more attractive. So may be the case of switching lanes on the road when the traffic on the driver's road moves slowly. Note, unlike the envelopes problem, which calls for a one-shot decision (as are many other problems in experimental decision making) in a completely symmetric, information-impoverished state, behavior facing real-world switching dilemmas may differ, in part because they considerably depend, for better or worse, on experience with repeated decisions, at least more than they do on considerations such as loss aversion (see Erev, Ert & Yechiam, 2008; Rakow & Newell, 2010; cf Erev et al, 2010): People appear to follow a heuristic subjectively perceived consonant with what they might reason from experience. This seems to be the behavior of participants in the study reported here.

Even more cogently, facing a challenge to perform especially well, as well as a chance to mold on the fly a strategy that apparently could be tested against continual feedback, many participants in this study were tempted to better their lot by switching frequently when they found themselves in the slower moving lane. People presumably do not do it that often in real-world situations, but that may be due more to inherent costs (for one, risk of collision, likelihood of being stopped by police on charge of dangerous weaving) than to discretion and sober judgment.[8]

Thus, switch seeking may be fairly prevalent, quite probably more than commonly assumed, typically when the person perceives her state as worse than the alternative one, and believes, often erroneusly, to have some clue that it would continue to be worse if she did not switch.

How general could these findings be? Evidently, they have been observed in an experimental situation, furthermore one in which the task was to optimize performance in a particular decision problem in which stakes were not high. On the other hand, that task was meant to simulate an everyday activity in which switching behavior seems to be fairly common. It is, thus, not far-fetched that the task does represent fairly well a mundane problem, and the findings do demonstrate a tendency that is considerably potent in real life.

Another interesting question is whether individuals who frequently switch lanes on the road or trade their stocks for other ones are more inclined to also seek more significant changes in life (such as changing working places, occupations, spouses). Answering that surely requires much further research.

Implications for normative decision making

Is decision making behavior, in the illustrative instance presented above or in others, really determined by reasoning of the type hypothesized here?

People who appear to follow a rule of behavior clearly do not have to be aware of following it, let alone to have deliberately formed it. If people do use the kind of reasoning we hypothesize they do, it must be by and large in a pre-rational stage, presumably by what is now customarily called *system-1* (see Kahneman, 2003; Sloman, 1996; Stanovich & West, 2000).

Anyhow, the pitfall is surely there. The mere possibility that people sometimes fall in it, or even just tempted by it, calls for some rethinking not only of the intuitions underlying much of human real-world decision making, but also of the *normative* rule of *indiscriminately* choosing between options by expected gain or expected utility (or by any other egocentric decision rule; e.g., choosing by median squared gain or by maximal gain) in the way typically prescribed.

The pitfall seems to be due to a fundamental conceptual habit: Postulating that (a) by default one's own state is given[9], (b) her uncertainty is due just to the state of the outside world, (c) her decisions must strive to maximize some function of her possible outcomes (e.g., gains) in view of that uncertainty.

Though maximizing expected gain (or expected utility) is surely quite sensible in the case that customary postulation is valid, the problem is that uncertainty may sometimes apply to one's own state no less than to the world's state. For example, when one rides a slow bus, she may consider switching to another one based on what she perceives is the former's present rate of progress and what she estimates (or fancies) is the latter's rate of progress.

She rarely tries to rather formulate the problem at the outset from a less egocentric angle: Given that there are two buses with different rates of progress, what is the chance that I am on the faster moving one right now? Do I have at all any pertinent data to decide that? In case I do not have such data (say, because the other bus is not in sight), for what I know I am equally likely to be riding now the faster moving bus as I am to be riding the slower moving one. Hence, unless I somehow obtain data sufficient to tip the balance between the horns of my dilemma, I might as well stick with the status quo.

Thus, a key factor in motivating a switch of the sort studied here may be the egocentric disposition. Granted, in numerous states egocentric framing is as reasonable as can be. The trouble is that in some other cases it drives people to seemingly maximize expected gains even when no base reference point exists. The distinctive feature of such a case seems to be that the values carried by the two alternatives at decision time are either unknown or totally unpredictive of the future.

To dissuade people from behaving in that way in such cases, it seems advisable to practice acentric framing. That requires some sort of elective dissociation. A way to avoid egocentric errors of the kind discussed above is to exercise just that, elective dissociation, namely to try viewing any situation from an outside perspective: "There are players; I better start by alienatedly calling them A and B; then let us assume that A happens to be me. What I should do is what A should have done from the alienated, outside perspective. Yet, from that perspective, A should not do anything different than B does".

That seems quite different from maximizing expected value (or for that matter, any other positive function) of gains as if the worth of one's state (portfolio, health, time-to-exit-a-queue, job, career) was definite.

Thus, the time-honored, normative rule of deciding between options by expected monetary value of possible outcomes (Bernoulli, 1954/1738), or by subjective expected utility (von Neumann & Morgenstern, 1944; cf earlier concepts of utility such as Gabriel Cramer's "moral expectation" concept discussed in Bernoulli, 1954/1738, or the one proposed by Jevons, 1965/1871), may be less universally adequate than customarily postulated. A caveat is called for: The rule should not be applied uncritically, rather only when amounts of gains and losses were either explicitly known or calculable from other data (e.g., the ratio between a consequence state and a known status-quo state).

One might ponder how frequently the latter condition is met. A tacit assumption underlying that rule is that humans aspire to maximize gains (typically, the subjective values of gains), namely increases with respect to the status-quo, or to some other reference point, whatever it might be. It is well established that in many cases, persumably most cases, they do so (e.g., Kahneman & Tversky, 1979), but it is still hard to repudiate the stance that they would sometimes do better to maximize their total assets (in the broad sense, including not just property but also health, job, etc), and typically the subjective utility of total assets. There is some evidence that they sometimes tend to do that (Kahneman, 2003). Yet, how often do people definitely know both their status-quo state of assets and their utility function in personal assets? Very rarely, so it seems, as Jevons (1965/1871) noted almost a century-and-a-half ago (ibid, p. 52).

And even in the case that people maximize subjective value of outcomes, how often do people definitely know those values? They probably do take some subjective magnitudes regularly into account in some *implicit* computation that their brain does while weighing alternatives intuitively, but can they normally retrieve them to be explicitly considered when they are to apply a normative rule in a deliberate fashion as sometimes required?

It, thus, seems that we often make decisions in a fuzzier state of knowledge than we typically realize. In that case, trying the alienated, acentric heuristic suggested above may be advisable.

References

Arkes, H.R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*, 124-140.

Avrahami, J., & Kareev, Y. (2011). The role of impulses in shaping decisions. *Journal of Behavioral Decision Making*, *24*, 515-529.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research*, *30*, 961-981.

Bell, D. E. (1988). Diappointment in decision making under uncertainty. In D.E. Bell, H. Raiffa and A. Tversky. (Eds.) *Decision Making: Descriptive, normative, and prescriptive interactions* (pp. 358-383). Cambridge: Cambrifge University Press.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica, 22*, 23-36. (Translated from Italian; originally published as "Specimen theoriae novae de mensura sortis" in 1738).

Brams, S., & Kilgour, D. (1995). The box problem: To switch or not to switch. *Mathematics Magazine, 68*(1), 27-34.

Butler, S.F., & Nickerson, R.S. (2008). Keep or trade? An experimental study of the exchange paradox. *Thinking and Reasoning, 14*, 365-394.

Chapman, L.J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior, 6*, 151-155.

Christensen, R. & Utts, J. (1992). Bayesian resolution of the "exchange paradox". *The American Statistician, 46*, 274-276.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, 112*, 912-931.

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making*, *21*, 575-597.

Erev, I., Ert, E., Roth, A.E., Haruvy, E., Herzog, S.M., Hau, R., Hertwig, R., Stewart, T., West, R., & Leibiere, C. (2010). Choice prediction competition: Choices from experience and from Description. *Journal of Behavioral Decision Making*, 23, 15–47.

Fishburn, P.C. (1983). Transitive measurable utility. *Journal of Economic Theory, 31,* 293-317.

Garland, H. (1990). Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Psychology*, *75*, 728-731.

Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: free Press.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*, 295-314.

Griffin, D., & Kahneman, D. (Eds.) (2002). *Heuristics and biases*. New York: Cambridge University Press.

Granberg, D., & Brown (1995). The Monty Hall dilemma. *Personality and Social Psychology Bulletin*, *21*, 711-723.

Hughen, J.C., & McDonald, C.G. (2005). Who are the noise traders? *The Journal of Financial Research, 28*, 281-298.

Jevons, W.S. (1965). *The theory of political economy* (reprinted 5th ed.). New York: A.M. Kelley. (Originally published in 1871).

Kahneman, D. (2003). A perspective on judgment and choice. *American Psychologist, 58*, 697-720.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 111-132.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39,* 341-350.

Kareev, Y. (2000). Seven (and indeed, plus or minus two) and the detection of correlations. *Psychological Review, 107*, 397-402.

Karmarkar, U.S. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance, 21*, 61-72.

Kileen, P. (1994). Mathematical principles of reinforcement: Based on the correlation of behaviour with incentives in short-term memory. *Behavioral and Brain Sciences*, *17*, 105-172.

Kubovy, M., & Gilden, D. L. (1990). Apparent randomness is not always the complement of apparent order. In G. Lockhead & J. R. Pomerantz (Eds.), *The perception of structure: Essays in honor of Wendell R. Garner* (pp. 115-127). Washington, DC: American Psychological Association.

Lopes, L. L., & Oden, G. C. (1987). Distinguishing between random and nonrandom events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 392–400.

Machina, M.J. (1982). Expected utility analysis without the independence axiom. *Econometrica, 50*, 277-323.

Markowitz, H. (1952). The utility of wealth. *Journal of political economy, 60*, 151-158.

Nalebuff, B. (1989). Puzzles: The other person's envelope is always greener. *Journal of Economic Perspectives, 3*, 171-181.

Nickerson, R.S. (2002). The production and perception of randomness. *Psychological Review, 109*, 330-357.

Nickerson, R.S., & Falk, R. (2006). The exchange paradox: Probabilistic and cognitive analysis of a psychological conundrum. *Thinking and Reasoning, 12*, 181-213.

Petrocelli, J.V., & Harris, A.K. (2011). Learning inhibition in the Monty Hall problem: The role of dysfunctional counterfactual prescription. *Personality and Social Psychology Bulletin*, *37,* 1297-1311

Rakow, T., & Newell, B.R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making, 23*, 1–14.

Redelmeier, D.A., & Tibshirani, R.J. (1999). Why cars in the next lane seem to go faster? *Nature, 401*, 35.

Redelmeier, D.A., & Tibshirani, R.J. (2000). Are those drivers really going faster? *Chance, 13*, 8-14.

Samet, D., Samet, I. & Schmeidler, D. (2004). One observation behind two-envelope puzzle. *American Mathematical Monthly, 111*, 347-351.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, *1*, 7-59.

Savage, L.J. (1954). Historical and critical comments on utility. In L.J. Savage*, The foundations of statistics* (pp. 91-104). New York: Wiley.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3-22.

Sobel, J. (1994). Two envelopes. *Theory and Decision, 36*, 69-96.

Stanovich, K.E., & West, R.F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences, 23*, 645-665.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 455-458.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Zabell, S.(1988). Symmetry and its discontents. In B. Skyrms and W.L. Harper (Eds.), *Causation, chance, and credence* (pp. 155-190). Dordrecht : Kluwer Academic Publishers.

Footnotes

1. For example, Nickerson and Falk (2006) propose the following solution for the version of the paradox featuring complete uncertainty with respect to envelope and contents: One envelope contains X dollars; the expected value of the contents of the other envelope is 1.25X dollars. Since I do not know which envelope I hold, the expected value of its contents is 1.125X dollars, and since that exactly is also the expected value of the contents of the other player's envelope, both of us have no reason to trade. Thus, the desired conclusion is reached.

2. That the customary formulation arouses the fallacy does not entail that it is impossible to obtain a rational solution within it. So is the solution proposed by Nickerson and Falk, described in Footnote 1. Note, however, that the line of reasoning used in that solution is somewhat awkward, since its first step asserts an unnecessary asymmetry between envelope contents (one envelope contains X dollars; the expected value of the contents of the other one is 1.25X dollars). The acentric formulation does away with any asymmetry. It is, therefore, more apt for avoiding fallacies.

3. Note that *any* deviation from acentricity leads to a bias (Jeff Miller, personal communication). Suppose I take an *altercentric* point of view, assuming that the other player has a definite amount, Y, while I have either 2Y dollars or Y/2 dollars. In that case, my choice is between having Y in case I opt for trading, and having an even mixture of two states of my own assets, 2Y or Y/2, in case I opt not to trade. Since the expected value of the latter is 1.25Y dollars, I would be tempted to strictly prefer it (rather than be indifferent between the options, as the acentric point of view leads to).

4. *Not* $0.5 \cdot 2c + 0.5 \cdot c2 = 1.25c$, as suggested by an egocentrically disposed intuition (at least of one expert reader of a previous version of this paper).

5. Actually, such a conclusion from any conceivable base of evidence is quite absurd, since residual life expectancy, being sensitive to multiple factors as well as to present age, is not likely to be affected by any single T in such a simple manner. That, however, is no guarantee that people, and not only laypersons, do not tend to draw such conclusions.

6. Ironically, that conclusion is false, to the extent that the preference for higher rate of progress is due to a motivation to shorten one’s travel time. Though people seem to rarely realize this, it is trivially true that since time and rate are reciprocally related, the effect on travel time of reducing one’s progress rate by a factor of R is absolutely *larger* than the effect of increasing it by the same factor.

7. In a pre-test designed to get some idea of the prevalence of a disposition to switch, we asked 58 students to imagine themselves in the following situation: "You are last in a queue for a cashier in a

small supermarket, and the pace of your progress seems to be slow. There is another cashier in the supermarket. You cannot see its queue, but you are told that the two queues are of the same length. You are further told that the pace of progress in one of the two queues is twice as fast as in the other one, but you are not told which of them is the faster going one".
Asked whether they would choose to switch to the other queue, as many as 36% of the respondents replied positively. This seems to be a fairly high proportion, considering that there are quite a few reasons to stick with the status-quo - some rational (e.g., "It takes extra time to leave the present queue, then move to the other one"), some irrational (e.g., "With my bad luck, better not pushing it"). On top of that, when presented in a further question with the mathematical argument that underlies the exchange paradox, 29% of the respondents found that argument reasonable.

8. Another possible cause is that drivers may perceive the switch-stay dilemma somewhat differently. Participants in the experiments reported here were explicitly told that speeds in the two lanes were always related by either a given ratio (1/2) or its inverse (2). That feature of multiplicative symmetry characterizes also the envelopes problem. The paradox results from the fact that such multiplicative symmetry probably lures the participant by appearing to preclude additive symmetry, namely by suggesting that the gain exceeds the absolute value of the cost. If drivers on a real-world road perceived the dilemma as having rather additive symmetry (namely, gain equals loss, in absolute values), they would have no temptation to overswitch. It does not seem very plausible, though, that people perceive the possible relationship between two speeds, or of any other two rates of a continual process for that matter, in terms of a difference (bewteen them or between their distances from the status-quo). Perceiving it in terms of ratios seems to be the natural way.

9. As argued in the section "Source of the fallacy", even when one realizes that her state is unspecified, as it is in the envelopes problem, she seems to regard it as definite (for example, "I have Y dollars in my envelope"), thereby relegating the uncertainty to the state of the outside world (in this particular instance, the content of the other player's envelope – either 2Y or Y/2). A conceivable way such an idea may come about is by tacitly presuming an asymmetric narrative of the generating process: I won Y by some allotment or lottetry; then the other player was endowed either 2Y or Y/2 by a flip of a fair coin.

Acknowledgments

We are indebted to Ziziana Lazar for programming the experiments, to Amit Pomeranz for assistance in developing software for various purposes, Ma'ayan Ben-Eliezer, Aviv Ben Shabat, Elad Gvirtz and Nadav Hirsch for running the experimen, and to Meira Ben-Gad for helping in editing and proofreading. We also wish to thank two anonymous reviewers, as well as the editor, Marco Bertamini, for useful comments and advices.

Appendix A

Do you feel you switched lanes frequently?
By your estimate, how many times did you switch?

[new page]

Have you used to switch in a certain state that you can describe?
If so, what was it?

[new page]

When you opted for switch, what was your main argument for doing so?

[new page]

Since you must have been aware of the risk that the switch would harm you, why did you nonetheless opted for it when you did?

[new page]

Now imagine yourself in the following situation:
You and some other person receive each a sealed envelope. In each envelope there is a sum of money. You are not told what the sums are, but you are informed that one of the sums, which one is unbeknown to you, is twice as high as the sum in the other envelope.
You are forbidden to open the envelope for an hour. However, you are allowed to exchange envelopes with the other person.

Would you choose to exchange envelopes or to keep the envelope you received? (think about that for at least a minute at least before you reply).

[new page]

In any event, explain your choice fully.

[new page]

Some people choose to exchange envelopes based on the following argument: The probability is 1/2 that the other person received twice as much as I did, and 1/2 – that he received half of what I received. So, if I exchange envelopes with him, my chance to earn a sum matching what I have in my envelope equals my chance to lose a half of what I have there. Hence, the exchange is worthwhile.

In view of this argument, would you choose to exchange envelopes or to keep the envelope you received? (Give it some thought, for at least a minute, before you reply)

[new page]

In any event, explain your choice fully

Table 1. Mean correlations between response (switch=1, stay=0), as a function of trial history, and (a) actual speed of traffic in the participant's lane, (b) relative speed between lanes.

	Variable being correlated with response		
	Actual speed	Relative speed	Mean difference
Trial history			
ARSS-encouraging (N=62)	-0.295	-0.478	-0.183
ARSS-discouraging (N=32)	-0.196	-0.244	-0.048
Overall (N=94)	-0.262	-0.399	-0.137

Table 2. Mean per-participant switch rate (standard deviation in parentheses) as a function of trial history, participant lane speed and sub-session (across all participants - in A; across just the participants in the "others" category - in B).

		Participant lane speed		
	Sub-session	Faster	Slower	Overall
ARSS-encouraging (N=62)	First half	0.190 (0.206)	0.690 (0.289)	0.441 (0.354)
	Second half	0.206 (0.259)	0.663 (0.307)	0.435 (0.364)
	Overall	0.198 (0.233)	0.677 (0.297)	
	Sub-session	Faster	Slower	Overall
ARSS-discouraging (N=32)	First half	0.273 (0.226)	0.517 (0.306)	0.395 (0.294)
	Second half	0.285 (0.290)	0.500 (0.307)	0.392 (0.315)
	Overall	0.279 (0.258)	0.508 (0.304)	

A

Participant lane speed

	Sub-session	Faster	Slower	Overall
ARSS-encouraging	First half	0.297 (0.197)	0.501 (0.203)	0.381 (0.239)
(N=38)	Second half	0.333 (0.259)	0.455 (0.199)	0.383 (0.266)
	Overall	0.315 (0.229)	0.478 (0.201)	

	Sub-session	Faster	Slower	Overall
ARSS-discouraging	First half	0.323 (0.221)	0.438 (0.247)	0.399 (0.224)
(N=26)	Second half	0.346 (0.290)	0.420 (0.240)	0.394 (0.237)
	Overall	0.335 (0.255)	0.429 (0.242)	

B

Table 3. Mean per-participant ARSS-compatible rate (standard deviation in parentheses), as a function of trial history, participant lane speed and sub-session (across all participants - in A; across just the participants in the "others" category - in B).

		Participant lane speed		
	Sub-session	Faster	Slower	Overall
ARSS-encouraging (N=62)	First half	0.810 (0.206)	0.691 (0.289)	0.750 (0.257)
	Second half	0.794 (0.259)	0.663 (0.307)	0.729 (0.290)
	Overall	0.802 (0.233)	0.677 (0.297)	
	Sub-session	Faster	Slower	Overall
ARSS-discouraging (N=32)	First half	0.727 (0.226)	0.517 (0.306)	0.622 (0.287)
	Second half	0.715 (0.290)	0.500 (0.307)	0.608 (0.315)
	Overall	0.721 (0.258)	0.508 (0.304)	

A

Participant lane speed

	Sub-session	Faster	Slower	Overall
ARSS-encouraging (N=38)	First half	0.703 (0.197)	0.501 (0.203)	0.602 (0.223)
	Second half	0.666 (0.259)	0.455 (0.199)	0.561 (0.253)
	Overall	0.685 (0.229)	0.478 (0.201)	

	Sub-session	Faster	Slower	Overall
ARSS-discouraging (N=26)	First half	0.677 (0.221)	0.438(0.247)	0.557 (0.261)
	Second half	0.654 (0.290)	0.420 (0.240)	0.537 (0.289)
	Overall	0.665 (0.255)	0.429 (0.242)	

B

Table 4. Pearson correlation coefficients between per-participant differential reinforcement rate in trials 2-29 and three per-participant measures of performance in trials 30-60 (excluding 47 & 48), separately for the two different history conditions (in A), and pooled across conditions (in B) .

		Differential reinforcement rate in trials 2-29 in which speed in participant's lane is	
	Performance in trials 30-60	Slower	Faster
ARSS-encouraging (N=62)	Stay rate in "Faster" trials	0.401***	0.387**
	Switch rate in "Slower" trials	0.598***	0.536***
	ARSS-compatible rate	0.597***	0.529***
ARSS-discouraging (N=32)	Stay rate in "Faster" trials	0.053	0.092
	Switch rate in "Slower" trials	0.049	0.073
	ARSS-compatible rate	0.051	0.132

* p<0.05, ** p<0.01 , *** p<0.001

A

	Differential reinforcement rate in trials 2-29 in which speed in participant's lane is	
Performance in trials 30-60	Slower	Faster
Stay rate in "Faster" trials	0.308**	0.315**
Switch rate in "Slower" trials	0.463***	0.438***
ARSS-compatible rate	0.462***	0.445***

* p<0.05, ** p<0.01 , *** p<0.001

B

Figure captions

Figure 1. An illustration of the participants' screen. While the cars displayed were heading toward the top of the screen, the participant's car was never seen moving up; it was rather stationary at the bottom of the screen (in this case in the left lane) and was somewhat larger. The surroundings (marked by icons of electricity poles on the road margins) "moving" toward the bottom. All other cars within the same lane were also seen as basically stationary (except for slight jitters). Cars in the other lane, when visible, were seen "moving" up when their speed was higher or down when their speed was lower.