How initial confirmatory experience potentiates the detrimental influence of bad advice

Markus R. Staudinger *, Christian Büchel

Department of Systems Neuroscience, Center for Experimental Medicine, University Medical Center Hamburg–Eppendorf, Hamburg, Germany

A R T I C L E   I N F O

Article history:
Accepted 28 February 2013
Available online xxxx

Keywords:
Advice
FMRI
Priming effect
Reinforcement learning
Willingness to pay
Ventral pallidum

A B S T R A C T

In everyday life, expert advice has a great impact on individual decision making. Although often beneficial, advice may sometimes be misleading and cause people to pursue actions that entail suboptimal outcomes. This detrimental effect may diminish over time, when individuals have gathered sufficient contradicting evidence. Given the strong influence initial information has on opinion and personality impression formation, we aimed to investigate whether initial advice-confirmatory experience potentiates the rigidity with which persons stick to misleading advice. Furthermore, we intended to characterize the neuronal basis of such putative priming effect. While undergoing functional magnetic resonance imaging (fMRI), participants selected between probabilistically reinforced symbols and were given the misleading tip that two low-probability symbols had a high reinforcement probability. One of these symbols initially received manipulated advice-congruent positive feedback (PF), the other one advice-incongruent negative feedback. Behaviorally, participants were impaired at learning to avoid advice-receiving symbols and overvalued them in terms of willingness to pay (WTP) in an auction market. Crucially, initial PF potentiated all effects. Greater ventral pallidal response to initial but not later PF during learning predicted higher behavioral WTP. Our results demonstrate that the nature of the very first advice-related experience already determines how strongly misleading advice will influence learning and ensuing decision making—an effect that is mediated by the ventral pallidum. Thus, in contrast to conventional reinforcement learning, learning under the influence of advice is susceptible to primacy effects. The present findings advance our understanding of why false beliefs are particularly difficult to change once they have been reinforced.

© 2013 Published by Elsevier Inc.

Introduction

In our everyday life, we aim to identify actions that are followed by rewards. Learning about such environmental contingencies can come in different ways. We may learn via personal experience of reward and punishment (trial-and-error learning). However, in many instances, following advice is a faster and less costly strategy of determining preferable actions (Bandura, 1977; Biele et al., 2009; Engelmann et al., 2009). Such a strategy carries an inherent risk though. Sometimes advice—like instructions and rules—produces inflexible, rigid patterns of behavior that are insensitive to actual environmental contingencies (Hayes, 1993). For instance, after having received one-time misleading (i.e., bad) advice, agents adhered to recommended actions although they led to far less optimal outcomes than unexplored, alternative ones (Doll et al., 2009, 2011). Skinner suggested that advice has that effect because it specifies contingency expectations (Skinner, 1966), causing individuals to seek out and favor advice-confirming information, the so-called confirmation bias (Nickerson, 1998; Ploos, 1993), and to dismiss any evidence to the contrary.

Although bad advice can exert a long-lasting detrimental influence on learning, behavioral rigidity may diminish over time, when individuals have gathered sufficient advice-contradicting evidence. Does it play a role when the first contradicting evidence is coming in? More precisely, does the character of the very first advice-related experience—which could be advice-congruent or -incongruent—already determine how rigidly a person will stick to bad advice? Social psychology research has demonstrated that primary information disproportionately influences both opinion and personality impression formation (Anderson and Barrios, 1961; Crano, 1977; Cromwell, 1950; Jones et al., 1968)—an effect termed the law of primacy in persuasion. For instance, Lund (1925) observed that the first of two opposed arguments on a controversial topic is more effective in changing individuals’ attitudes. Stone (1969) in turn showed that in role play lawsuits primary testimonies affect jurors’ final verdicts to a greater degree than later testimonies. In view of this, primary experiential information could be expected to also influence how strongly bad advice impacts learning. Initial confirmatory experience might reinforce advice-induced a priori beliefs to such a degree that advice-following behavior becomes resistant to...
change. Initial advice-contradicting experience, however, might entail a
discounting of the misleading advice and a change in action policy.

Our aim in the present study was to investigate whether initial
advice-confirmatory experience potentiates the detrimental influence
of misleading advice on learning and decision making. Furthermore,
we intended to characterize the neuronal basis of such putative primacy
effect on advice. While undergoing fMRI, participants were repeatedly
presented with four fixed pairs of probabilistically rewarded Japanese
Hiragana symbols (Fig. 1A) (Frank et al., 2004). In each pair, one symbol
had a higher probability of being correct (60%) than the other (40%). Vol-
unteers had to learn to select the symbol with the higher probability via
trial-and-error. Critically, before entering the scanner, participants were
given the misleading tip that two of the four 40% symbols—symbols D
and H—possibly had the highest probabilities of being correct among
all symbols. To investigate the effect of initial advice-confirmatory ex-
perience, we manipulated the reinforcement feedback during the
first three trials in all four pairs. Crucially, while choosing the advice-
receiving symbol H was definitely correct during the first two trials
and incorrect in the third (initial advice-congruent condition), choosing
the other advice-receiving symbol D was incorrect in the first two trials
but correct in the third (initial advice-incongruent condition). The feed-
back for the no-advice 40% symbols B and F was manipulated in a similar
way so that they could serve as control symbols. After the third trial,
feedback became probabilistic. Subsequently, outside the scanner, par-
icipants completed a test task where they saw all possible combinations
of symbols and had to choose in the absence of reward feedback
(Fig. 1B). Afterwards, for each symbol, volunteers completed a Becker–
DeGroot–Marschack (BDM) auction (Becker et al., 1964) where they
bid on a lottery ticket whose probability to win €8 equaled the probabil-
ity specified by the displayed symbol (Fig. 1C). These BDM auctions en-
abled us to determine symbol-specific willingness to pay (WTP) values
(Chib et al., 2009; Plassmann et al., 2007). Finally, participants completed
a second fMRI session where they made purchase decisions for such lot-
tery tickets at predetermined prices.

We hypothesized that one-time misleading advice results in
suboptimal choice preference of advice-receiving 40% symbols D and
H over 60% symbols during learning and test (Doll et al., 2009,
2011) as well as in higher WTP values of D and H as compared to the
no-advice 40% symbols B and F. Critically, we expected early
positive reinforcement to potentiate these effects: initial, positive,
advice-congruent feedback should result in poorer performance on
H than on D throughout the experiment, but entail no performance
difference between control symbols F and B.

Previous studies have implicated the ventromedial prefrontal cor-
tex (vmPFC), putamen, caudate, nucleus accumbens (NAcc) (Biele et
al., 2011; Jocham et al., 2011; Li et al., 2010), and NAcc-downstream

![Fig. 1. Experimental design. (A) Probabilistic reinforcement learning fMRI session. Participants were repeatedly presented with four fixed pairs of probabilistically rewarded symbols (Hidden P: 60% vs. 40%) and had to learn to choose the symbol with the higher probability. Before entering the scanner, participants were given the misleading tip that symbols D and H possibly had the highest probabilities of being correct. Feedback during the first three trials was manipulated in all four pairs. Crucially, choosing H was definitely correct during the first two trials and incorrect in the third (initial advice-congruent feedback), choosing D was incorrect in the first two trials but correct in the third (initial advice-incongruent feedback). (B) Test task. Participants were presented with all possible combinations of symbols and had to choose in the absence of feedback. A high percentage of avoid decisions (preferring 60% symbols when being paired with 40% symbols) indicates successful learning. (C) BDM task. Participants bid on lottery tickets whose probabilities to win €8 equaled the objective learning phase probabilities of the displayed symbols, enabling us to measure symbol-specific WTP. (D) Lottery ticket auction fMRI session. Again, participants made purchase decisions for lottery tickets, only at predetermined prices (equal to the median WTP of all symbols ±€0.10), allowing us to determine neural correlates of WTP.](http://dx.doi.org/10.1016/j.neuroimage.2013.02.074)
ventral pallidum (VP) (Haber and Knutson, 2010; Ito and Doya, 2009) in reinforcement learning. Consequently, we suspected that if initial positive reinforcement increased the impact of misleading-advice, this behavioral effect would be mediated via initial feedback responses in one or more of the reinforcement feedback-sensitive areas.

Materials and methods

Participants

Thirty-five, right-handed, male volunteers participated in the study. Three volunteers failed the learning criteria in the reinforcement learning session; did not participate in any of the ensuing sessions, and were excluded from analyses. Participants were compensated with €10/h plus additional payment depending on their learning performance as well as on the outcome of the lottery session. All participants gave written informed consent. After the experiment, participants received detailed verbal and written information about the study objectives. The study was approved by the local ethics committee.

Tasks

Probabilistic reinforcement learning task (1st fMRI session)

While undergoing fMRI, participants completed a probabilistic reinforcement learning task (Doll et al., 2009; Frank et al., 2004) in which they selected between probabilistically rewarded Japanese Hiragana symbols (60% vs. 40%; Fig. 1A). Volunteers were instructed to learn to select the symbol with the higher probability via trial-and-error. Task instructions read:

"On the screen, pairs of two symbols will appear. The fixed combinations are... (all pairs were displayed). In each trial, you have to select one of the two symbols. Afterwards, you will be informed whether your choice was correct or incorrect. In a given trial, exactly one symbol will be correct, whereas the other one will be incorrect. Please note: NO symbol is correct ALL THE TIME. However, in each pair, one symbol is correct MORE OFTEN than its partner symbol. Your task is to find those symbols. In the beginning, you will not know which is which. Try to learn this over time and then pick the respective symbols. Please note also: when a symbol is correct in a given trial does not depend on which side (left or right) it appears (...) Training will stop after you have reached a certain learning criterion. Of all the training trials, we will draw 10 at random and reward you with 0.50€ each time your choice had been correct."

Furthermore, on the last page of the instruction sheet, participants read the misleading statement:

"Here is a tip: The following two symbols possibly have the highest probabilities of being correct. Try to memorize them!" ... (the Hiragana symbols for D and H were displayed in between-subject randomized order).

Subsequently, participants were asked to give a summary of the instructions in order to check whether they had understood the nature of their task. Furthermore, they indicated on the first page of the instruction sheet which symbols possibly had the highest probabilities of being correct (Doll et al., 2009). All participants managed to point out the symbols D and H on first attempt.

We manipulated the feedback during the first three trials in all four pairs (see Introduction). After the third trial, reward feedback became probabilistic. In each trial, participants had 3.5 s time to select and—after a variable period of 1–8 s—received feedback (2 s) whether their choice had been correct (“Correct!”) or incorrect (“Incorrect!”); misses fed back as “Too late!” Symbol location (left, right) was varied. Symbol-to-condition assignment varied between participants.

Test task (post-scan)

During the test phase (Fig. 1B), participants saw all possible 28 pair-wise symbol combinations 5 times. To test what they had learned, they had to choose in the absence of reward. Participants were instructed to choose that symbol which they believed to have had the higher probability of being correct during learning (allotted time: 4 s). Thus, in the absence of feedback, volunteers had to rely completely on what they had learned during the learning session.

BDM task (pre-scan)

Subsequently, for each symbol, participants completed one BDM auction (Fig. 1C) (Becker et al., 1964). For each lottery ticket auction, participants had a budget of €4 and could bid between €0 and €4. Each ticket offered the opportunity to win €8 at a probability specified by the displayed symbol. The probability equalized the probability with which it was correct during the learning phase. Thus, to determine how much money should be reasonably spent on a given ticket, participants had to rely on their estimates of the underlying learning phase probability. Evidently, the WTP can be expected to increase with the individual estimate of the symbol-related probability of winning. In the present auctions, the optimal bid for a 60% chance of winning would have been €2.40, whereas the maximum bid for a 40% chance of winning should not have exceeded €1.60. The BDM auctions thus enabled us to determine symbol-specific WTP values (Chib et al., 2009; Plassmann et al., 2007) and to test whether misleading advice would continue to impair judgment outside the original learning environment. Task instructions are given in the Supplementary methods.

Lottery ticket auctions (2nd fMRI session)

Directly after the BDM auctions, participants completed a second fMRI session, during which they made purchase decisions for symbol-related lottery tickets at predetermined prices (Fig. 1D; 140 auctions in total, 15 per symbol). Prices were set equal to the median bid over all symbols made in the BDM auction, ± a variation of €0.10 (Chib et al., 2009). Again, each ticket offered the opportunity to play a lottery in which participants could win €8 at a probability specified by the displayed symbol. Volunteers had 4 s time to decide whether they would accept the offer, that is, pay the price and get to play the lottery if it was drawn in the post-scanning phase (accept: left button press; reject: right button press). Each choice was followed by a 1–8 s jittered period of a fixation cross display. After scanning, one auction was drawn from the pool of pre- and within-scanning auctions and its outcome was implemented (for a detailed description see Supplementary methods).

fMRI acquisition

fMRI data were collected using a 3-Tesla whole-body MRI system equipped with a 32-channel head coil (Trio; Siemens Medical Solutions). T2*-weighted functional images for both fMRI sessions were obtained using echo planar imaging (EPI; repetition time = 2620 ms, echo time = 25 ms, flip angle = 80°). Each volume comprised 44 slices with a voxel size of 2 × 2 × 2 mm and a gap of 1 mm. Slices were positioned parallel to the anterior–posterior commissure line. The field of view covered the whole brain, except for the most dorsal parts of the parietal lobes as well as for the most caudal aspects of the cerebellum.

fMRI preprocessing

Individual functional images were corrected for motion artifacts by realignment to the first volume of the run. Functional images were spatially normalized (2 × 2 × 2 mm) to an EPI template in the
Montreal Neurological Institute (MNI) space and spatially smoothed with a 6 mm full width at half maximum isotropic Gaussian kernel.

fMRI data analysis

First level analyses were performed to estimate parameters for the different conditions using general linear models. Intrinsic autocorrelations were accounted for by the first-order autoregressive AR(1) process, and low-frequency drifts were removed via high-pass filtering (128 s). All but the realignment regressors were modeled with their respective durations and were convolved with the hemodynamic response function.

Reinforcement learning session

We classified the first two, manipulated feedback trials per pair on the basis of the symbol actually chosen. The obtained categories were modeled as separate feedback regressors in the first level design, giving separate columns for initial correct A feedback, initial incorrect B feedback, initial incorrect C feedback, and initial correct D feedback. Later feedback trials were likewise divided into symbol-wise correct versus incorrect feedbacks (i.e., later correct A feedback, later incorrect C feedback, later correct G feedback, later incorrect G feedback). Furthermore, we introduced four regressors for the choice period depending on the symbol pair being displayed (i.e., A-B pair, C-D pair, etc.) and one nuisance regressor for missed choices. To investigate which neural region mediates the effect of initial advice-congruent reinforcement feedback on behavioral WTP, we implemented a second level multiple regression analysis. Here, individual WTP values for the symbol H were regressed against the contrast image for initial correct H feedback. Subsequently, we set up a one-way within-subject ANOVA with no-sphericity correction. We included contrast images for both initial and later reinforcement feedbacks for each 40% symbol, yielding one factor symbol with eight factor levels—initial and later incorrect B feedback, initial and later correct F feedback, initial and later incorrect D feedback, and initial and later correct H feedback—plus 32 subject regressors. Individual, symbol-related WTP values were included as covariates, and defined as interacting with the factor of symbol, giving eight covariate columns. This design allowed us to directly test for activation-WTP correlation differences between symbols and phases (initial, later).

Lottery ticket auction session

In the first level design, we introduced one regressor for the choice period, where the event was modeled with the actual choice duration. Missed choices were modeled using a separate nuisance regressor. Furthermore, we included two linear parametric modulators of the choice event: (a) the pre-scan WTP of the displayed symbol and (b) a second modulator encoding the actual choice, that is, a+1 for accepted offers, and a 0 for rejected offers (Chib et al., 2009).

Statistical inference

We report activations that survive family-wise error (FWE) voxel-level correction for multiple comparisons at $P < 0.05$. In cases of a priori regions of interest (ROI), FWE correction refers to small volume correction (SVC). SVC for a priori regions was based on masks of the cortical and subcortical Harvard-Oxford atlases as provided by FSL (FMRI Software Library), vmPFC activations were identified via the medial frontal cortex mask (Hare et al., 2011; Jocham et al., 2011; Wunderlich et al., 2009). The striatum mask comprised the putamen, caudate, and nucleus accumbens (Doll et al., 2009; Jocham et al., 2011; Li et al., 2010). For the VP, we modified the pallidum mask such that the search volume was restricted to MNI z < 0 (Pessiglione et al., 2007).

Results

Behavioral results

Probabilistic reinforcement learning task (fMRI scan)

Participants were instructed to learn to choose the better symbol in each pair, that is, that symbol that led to more frequent positive reinforcement (Doll et al., 2009; Frank et al., 2004, 2006; Jocham et al., 2011; Klein et al., 2007). We defined a learning criterion that required 308 volunteers to select the no-advice-pair 60% symbols A and E more than 50% in a given learning block of 80 trials (i.e., 20 trials per pair) (Doll et al., 2009; Frank et al., 2004, 2006). No learning criterion applied to the advice-pair 60% symbols C and G. Thirty-two out of 35 volunteers (mean age 27.6) met the learning criterion without exceeding the cut-off of 4 learning blocks (mean 1.53 ± 0.11 standard error). Moreover, preference for A and E during the last 10 learning trials was almost total (Fig. 2A).

To investigate learning over time, we calculated an ANOVA on the percentages of optimal 60% symbol choices (Fig. 2A) with the factors of advice (no advice: A; E; misleading advice: C, G), initial feedback manipulation (IFM; positive feedback: A, C; negative feedback: E, G), and time (first 10 vs. last 10 trials, please refer to Supplementary Fig. 1 for the whole time course of learning during the first block). We observed main effects of time ($F_{1,31} = 36.8, P < 0.001$), advice ($F_{1,31} = 98.8, P < 0.001$), and IFM ($F_{1,31} = 27.5, P < 0.001$) as well as significant interactions time × advice ($F_{1,31} = 8.5, P = 0.007$), time × IFM ($F_{1,31} = 11.4, P = 0.002$), and time × advice × IFM ($F_{1,31} = 8.2, P = 0.007$). The interaction time × advice shows that in no-advice pairs, participants easily learned to prefer A over B and E over F in the course of the learning phase. However, in misleading-advice pairs, participants did not learn to prefer C over D or G over H. Crucially and in line with our a priori hypothesis, the interaction time × advice × IFM demonstrates that the initial reinforcement manipulation influenced learning in misleading-advice pairs only. In no-advice pairs, participants started off with a high performance on the initially positively reinforced 60% symbol A, but performed poorly on the initially negatively feedback back 60% symbol E. Nevertheless, such a performance difference decreased over time so that during the last 10 trials, choice preference for E did not differ from A ($t_{31} = -0.5, P = 0.64$). In misleading-advice pairs, participants also began with a poorer performance on the initially negative feedback-receiving 60% symbol G (G vs. C during the first 10 trials). However, unlike in no-advice pairs, this difference was still present during late learning (G vs. C choices during the last 10 trials: $t_{31} = -2.3, P < 0.05$). Thus, initial advice-congruent, positive feedback on H as compared to initial advice-incongruent, negative feedback on D had caused volunteers to develop a stronger preference of H over G than of D over C and thus increased the effect of misleading advice.

Test task (post-scan)

In the absence of feedback, a crucial index of successful learning is the percentage of 40%-symbol-avoid-decisions (Doll et al., 2009), that is, decisions to prefer 60% over 40% symbols (Fig. 1B). A respective ANOVA on 40%-symbol-avoid-decisions with the factors of advice (no advice: B; F; misleading advice: D, H) and IFM (positive: F, H, negative: B, D) revealed a main effect of advice ($F_{2,62} = 46.2, P < 0.001$), and an interaction advice × IFM ($F_{2,62} = 4.5, P < 0.05$). Exploratory t-tests proved that both B and F avoid performances were way above 50% chance level (B: $t_{31} = 5.6, P < 0.001$, F: $t_{31} = 9.1, P < 0.001$), whereas D and H performances did not differ from chance (D: $t_{31} = 0.4, P = 0.70$, H: $t_{31} = -1.7, P = 0.11$). This demonstrates that the volunteers successfully avoided 40% symbols as long as they had not received pre-learning misleading advice. Again, the interaction advice × IFM confirms that the initial reinforcement manipulation influenced performance in advice-receiving symbols only. Initial two-time positive...

Please cite this article as: Staudinger, M.R., Büchel, C., How initial confirmatory experience potentiates the detrimental influence of bad advice, Neuroimage (2013), http://dx.doi.org/10.1016/j.neuroimage.2013.02.074
versus negative feedback did not alter F as compared to B, but decreased H as compared to D avoid performance.

### BDM task (pre-scan)

Participants properly valued the 60% symbols A and E higher than their former 40% partners B and F (Fig. 2C; $t_{31} = 7.30, P < 0.001$). Such 60% > 40% valuation was absent in advice-receiving pairs ($t_{31} = -1.27, P = 0.21$). We applied an ANOVA to 40% symbol WTP values (Fig. 2C) with the factors of advice (no advice, misleading advice) and IFM (positive; negative). It revealed a main effect of advice ($F_{1,31} = 33.8, P < 0.001$): as expected, misleading advice led to higher WTP valuation of advice-receiving (D and H) as compared to no-advice (B and F) 40% symbols ($t_{31} = 5.81, P < 0.001$). The ANOVA furthermore yielded a trend toward significance for the interaction advice × IFM ($F_{1,31} = 3.7, P = 0.06$). Because of our a priori interaction hypothesis, we tested a directed $t$-test that revealed a significant interaction (Fig. 2C; H–D > F–B: $t_{31} = 1.95, P < 0.05$). Thus, in accordance with the test results, initial positive reinforcement did not influence later WTP valuation in no-advice 40% symbols—WTP for F did not differ from B. However, it led to a difference in advice-receiving 40% symbols, where H was valued higher than D.

### Lottery ticket auction (fMRI scan)

In Fig. 2D, we plot percentages of within-scanner choices as a function of pre-scan WTP. The graph demonstrates that the likelihood of accepting an offer increased with the symbol-related WTP. However, this increase was not significant, *P* < 0.05.

Fig. 2. Behavioral results. (A) Reinforcement learning task. In no-advice pairs AB and EF, participants easily learned to prefer the 60% symbols A and E. In misleading-advice pairs CD and GH, participants failed to learn to prefer the 60% symbols C and G. Crucially, the initial feedback manipulation influenced learning in misleading-advice pairs only. Initial correct (A and C) versus incorrect (E and G) feedback did not entail differences in choice between A and E during the last 10 trials, but led to more frequent choices of C compared to G, speaking for a potentiation of the misleading advice effect via initial advice-congruent feedback. (B) Test task. Participants successfully avoided the no-advice 40% symbols B and F when being paired with 60% symbols, but were significantly worse at avoiding the advice-receiving symbols D and H. Initial correct versus incorrect feedback did not alter F versus B, but decreased H versus D avoid accuracy (indexed two-way interaction). (C) BDM results. Participants overvalued misleading-advice as compared to no-advice 40% symbols in terms of WTP. Furthermore, initial correct versus incorrect feedback did not alter F versus B WTP, but led to higher WTP for H than for D (indexed two-way interaction). (D) During the within-scanner lottery ticket auctions, the likelihood of accepting an offer increased with the symbol-related WTP. n.s. not significant, *P* < 0.05.
Our main goal was to investigate which neural region mediates the reinforcing effect of initial advice-congruent feedback on misleading advice. To this end, we tested whether neural responses to H-related reinforcement within the first two GH trials would predict H-related behavior at the latest experimental stage, that is, during WTP valuation. Individual WTP values for the symbol H were regressed against the contrast image for initial correct H feedback. We observed a positive correlation in the left VP (Fig. 3A; MNI coordinates $x = -10, y = -4, z = -8, t = 5.69; P < 0.05$ corrected). No significant correlation was found in the vmPFC or striatum. The scatter plot in Fig. 3A—extracted from the activation maximum—illustrates that higher individual VP response to initial correct feedback on H predicted greater, later WTP for H. We then tested whether neural responses were already associated with performance for H during earlier experimental stages, that is, during late learning and during test. Regressing the respective covariates against the contrast image for initial correct H feedback, we observed correlations in the left and right VP (% H choices during the last 10 training trials: $x = -10, y = -4, z = -8, t = 5.81; x = 10, y = 0, z = -12, t = 4.17; \%$ accuracy of avoiding H during test: $x = -10, y = -2, z = -10, t = 5.30; \text{all } P < 0.05$ corrected). Thus, higher individual VP response to initial positive reinforcement for H predicted greater individual percentages of H choices during the last 10 trials and less accuracy of avoiding H during test.

We further investigated whether the capacity of positive reinforcement activity to predict later WTP was restricted to early reinforcements and advice-receiving symbols. To demonstrate such exclusivity, initial positive reinforcement activations should be more positively correlated with H-related WTP than later positive reinforcement activations. At the same time, and to rule out a general primacy effect, initial positive reinforcement activations should be more positively correlated with WTP for H than WTP for F—H’s control symbol. We set up a one-way within-subject ANOVA analysis including initial and later feedback contrast images for all four 40% symbols. Individual symbol-wise WTP values were inserted as covariates, interacting with the feedback regressors (see Material and methods). This design allowed us to simultaneously test for the abovementioned differences.

---

**Fig. 3.** fMRI results reinforcement learning session. The VP mediates the reinforcing effect of initial advice-congruent feedback on misleading advice. (A) fMRI response to correct feedback for the symbol H during the first two GH trials correlated positively with H-related WTP valuation at the end of the experiment ($P < 0.05$ corrected). (B) Regression slopes of individual WTP values for all 40% symbols against corresponding initial (first two) and later feedback responses in the VP. Conjunction analysis: Initial correct feedback responses for H were more positively correlated with H-related WTP than later correct feedback responses. At the same time, initial correct feedback responses were more positively correlated with WTP for H than WTP for F—H’s control symbol. $P < 0.05$ corrected. Coordinates refer to MNI space.

---

Please cite this article as: Staudinger, M.R., Büchel, C., How initial confirmatory experience potentiates the detrimental influence of bad advice, NeuroImage (2013), http://dx.doi.org/10.1016/j.neuroimage.2013.02.074
in brain-WTP correlation. A conjunction analysis on the respective co-
variates (correlation between initial correct feedback and WTP for H > correlation between later correct feedback and WTP for H + correlation between initial correct feedback and WTP for H > correlation between initial correct feedback and WTP for F) (Nichols et al., 2005)
again revealed activity in the left VP (Fig. 3B; x = −12, y = −2,
z = −12, t = 3.81; P < 0.05 corrected). Again, no significant activa-
tions were found in the vmPFC or striatum. In Fig. 3B, we plot beta esti-
mates (i.e., slopes) for the regression of the symbol-wise WTP against the corresponding initial and later feedback activity, extracted from the VP maximum. The plots reflect the conjunction result: initial correct feedback responses in the VP predicted higher WTP for H, but later correct feedback responses did not. Furthermore, initial correct feedback responses correlated more positively with H-related WTP than with F-related WTP. This rules out the possibility of a generalized primacy effect of positive reinforcement in the VP. Finally, VP responses to initial incorrect feedback on D were not associated with later WTP val-
uation of D, showing that the VP was exclusively sensitive to initial (advice-congruent) positive reinforcement.

Lottery ticket auction session

The behavioral analysis had demonstrated that the scanner choices were in line with pre-scan WTP values. Thus, a single decision to accept or reject an offer most presumably involved a comparison between the symbol’s WTP and the depicted prize, which varied ±0.10 around the subject's median WTP. We therefore aimed to identify which of our a priori ROI encoded WTP during the scanner choices. Testing a one-sample t-test on the first level WTP parametric modulator (see Material and methods), we found that choice activation in the vmPFC correlated positively with pre-scan WTP (Fig. 4; t = 5.24; P < 0.05 corrected). This finding is consistent with its role in value-based decision making in general (Hare et al., 2009, 2011; Wunderlich et al., 2009) and with WTP in particular (Chib et al., 2009; Plassmann et al., 2007). No significant positive correlation was found in the striatum, hippocampus, or VP.

Discussion

Our data show that one-time misleading advice powerfully im-
pedes learning and decision making (Doll et al., 2009, 2011), an effect that even persists when individuals are placed outside the original learning environment. Most importantly, we can demonstrate that initial experiential information plays a decisive role in determining how rigidly misleading advice controls learning. Receiving manipu-
lated advice-confirmatory as opposed to advice-contradictory in-
formation early on in the learning process seems to strengthen false a priori beliefs about environmental contingencies. This causes participants to exhibit an increased behavioral rigidity—as evidenced by poorer behavioral performance for the initially positively reinforced instructed symbol H during learning, test, and the BDM auction.

Importantly, the character of the initial reinforcement experience does not influence learning in no-advice symbol pairs. Although par-
ticipants started off with a lower performance on the initially nega-
tively (E) than on the initially positively (A) fed back 60% symbol, they still learned to prefer the former to its suboptimal partner sym-
bol. It follows that—as is assumed in traditional reinforcement learn-
ing theories (Frank et al., 2007; Schultz et al., 1997; Sutton and Barto, 1988; Watkins and Dayan, 1992)—initial reinforcement does not play a role for the final learning success in conventional reinforcement learning (Frank et al., 2007). However, it is relevant during reinforce-
ment learning under the influence of advice, where specific weight is given to early feedback information. This finding concurs with primacy effects as observed in social psychology, where initially presented information tends to disproportionately influence the formation of opinions and personality impressions (Anderson and Barrios, 1961; Crano, 1977; Jones et al., 1968; Lund, 1925; Stone, 1969; Yates and Curley, 1986). Evidence from the domain of subjective probability revisi-
tion (Peterson and DuCharme, 1987) further corroborates our finding: When participants sample sequences of probabilistic, dichotomous events and repeatedly estimate the underlying probability distribution, early occurring events determine participants’ final estimates more than later ones.

Current theories posit that the tendency of humans to stick to suboptimal advice, instructions, and rules in spite of continuous nega-
tive outcomes is caused by a confirmation bias (Doll et al., 2011; Hayes, 1993; Nickerson, 1998; Skinner, 1966). Individuals seem to seek out and favor advice-confirming information, and dismiss contradicting evidence. Recent studies have raised the hope that the confirmation bias may be mitigated either directly via explicit instruc-
tions to consider alternative hypotheses/contradicting evidence or indi-
directly via stimulus material that renders alternative hypotheses/ 
contradicting evidence more salient (Ask and Granhag, 2005; 
Lehner et al., 2009; Lord et al., 1984). In line with the latter approach, we show that the confirmation bias resulting from misleading advice can be attenuated indirectly through repeated exposure of advice-
congruent evidence during early learning. Our behavioral data sug-
gest that during the early learning phase, advice-driven, false beliefs may be particularly susceptible to disconfirming information and therefore easier to be modified, resulting in a reduction of the con-
firmation bias during later learning.

At the neural level, we found that greater VP responses to initial pos-
tive reinforcement of misleading advice predicted worse behavioral performance throughout the experiment. Critically, this effect was re-
stricted to early and positive reinforcement and did not apply to no-advice symbols. Thus, the VP shows a sensitivity towards advice-
confirming information early during the reinforcement learning process. This finding concurs with its demonstrated involvement in reinforcement learning (Haber and Knutson, 2010; Ito and Doya, 2009). In fact, the VP is a central relay in the “affective” cortico-basal ganglia-thalamo-cortical loop that links the vmPFC, orbitofrontal cortex, and dorsal anterior cingulate with the NAcc and downstream VP, with the latter projecting back to the prefrontal cortex via mediodorsal thalamic nuclei (Alexander et al., 1990; Aron et al., 2009; Groenewegen and Trimble, 2007; Haber and Knutson, 2010). This circuit’s main func-
tion seems to lie in motivation and reward prediction, or, the learning of Pavlovian value (i.e., in the present experiment, symbol-outcome asso-
ciation) (O’Doherty, 2004; O’Doherty et al., 2004). Its contribution to instrument learning (i.e., in the present experiment, choice-outcome association) may be to increase the vigor or motivational loading of goal-directed action selection represented in the “associative” DLPC-caudate loop and vmPFC-caudate circuitry (Balleine and O’Doherty, 2010; Joel and Weiner, 1994; O’Doherty, 2004; Yin and

Fig. 4. fMRI results lottery ticket auction session. The vmPFC encoded WTP during scanner choices (P < 0.05 corrected). Coordinates refer to MNI space.

Please cite this article as: Staudinger, M.R., Büchel, C., How initial confirmatory experience potentiates the detrimental influence of bad advice, Neuroimage (2013), http://dx.doi.org/10.1016/j.neuroimage.2013.02.074.
Knowlton, 2006). Accordingly, the VP has been implicated in motivation and incentive salience (Berridge and Robinson, 2003; Mogenson et al., 1980; Pessiglione et al., 2007; Smith et al., 2009). Furthermore, from a cognitive psychology point of view, it has been argued that the confirmation bias is maintained via a combination of “hot” and “cold” cognitive processes (MacCoun, 1998). Whereas the latter may take place although individuals intend to be strictly objective, hot processes refer to—a—not necessarily conscious—motivation to find a priori beliefs confirmed. Thus, in the current experiment, “hot” processes might be reflected either in an overly optimistic a priori prediction of Pavlovian value (i.e., symbol-outcome association) (Doll et al., 2009) or in a high a priori motivation for advice confirmation. We suspect that the VP effects may be indicative of inter-individual differences in such a priori motivation or symbol start value. If the conditions were met that the initial feedback confirmed the advice—as was the case for the symbol H—feedback signals in the VP may have boosted symbol value and a priori beliefs and invigorated advice-following behavior to such a degree that it became resistant to change.

The imaging results further demonstrate that a brain region that mediates the reinforcing effect of initial confirmatory evidence on advice does not necessarily have to be categorically feedback-sensitive: Fig. 3A shows that nearly half of the participants deactivated in response to the initial confirmatory feedback. However, a significant response in reaction to the confirmatory evidence in the VP predicted poor choice behavior throughout the experiment. It is therefore valid to implicate the VP in the consolidation of advice-governed beliefs. Such consolidation could be driven by an interaction of top-down signals from the ventral striatum—which was feedback-sensitive throughout the experiment for all symbols—and some other unknown signal.

In sum, our data show that initial confirmatory experience is a potent reinforcer of misleading advice, rendering advice-following behavior resistant to change—an effect that is mediated by the VP. Thus, in contrast to conventional reinforcement learning, learning under the influence of advice is susceptible to primacy effects. The present findings can advance our understanding of why advice-governed behavior is particularly difficult to change once it has been reinforced.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.02.074.

Funding

This work was supported by the Landesexzellenzinitiative “neuroadt!” (MRS, CB) and DFG, SFB TRR 58 (B03) and SFB 936 (A06) to CB.

Acknowledgments

We thank Jan Peters, Jan Gläscher, Eszter Schoell, and Dharsnan Kumar for helpful comments.

References


As can advance our understanding of why advice-governed behavior is particularly difficult to change once it has been reinforced.


