Full title: A pupil-linked arousal mechanism for deciding to engage in future physical effort

Short title: Arousal and future effort

Authors

Irma T. Kurniawan[1*], Marcus Grueschow[1], & Christian C. Ruff[1]

Affiliations

[1] Zurich Center for Neuroeconomics
Department of Economics, University of Zurich
Blümlisalpstrasse 10
8006 Zürich

[*]Corresponding author (irma.kurniawan@gmail.com)

Abstract

An organism's behavioral success is determined by its ability to mobilize resources to overcome challenges. This ability involves the noradrenergic system, indicated by the finding that pupil-size increases proportionally with currently exerted effort. However, humans can deliberate in advance whether to engage in effort in the future. It remains unclear how effort is represented in such an anticipatory fashion during decision-making. We investigated this by measuring pupil responses while participants decided whether to accept or reject rewards that required effort execution after the experiment. We found a faster rate of pupillary dilation in decisions to accept high-effort rewards. This was accompanied by stronger fMRI activity in anterior cingulate cortex (ACC) and anterior insula: When accepting high-effort rewards, individuals with faster pupil dilation showed larger activity in these areas. Our results identify a brain process instantiating anticipatory arousal when humans prepare for a physical challenge, potentially reflecting simulated energization.

Introduction

Should I go to the gym tonight or should I skip training? The ability to select actions by considering their costs and benefits is crucial for survival in most animals [1]. Relatively unique to humans, however, is the remarkable ability to take such choices in a purely anticipatory fashion, deciding about potential future actions for which the potential benefits and costs are out of sight [2]. This ability is important for planning as it allows us to deliberate for sequences of actions whether the effort of overcoming all subsequent costs will be worth the associated rewards.

Indeed, humans constantly simulate future rewards to make decisions. Cues associated with reward trigger more vivid imagination of future events than neutral ones do [3]. There is also evidence that we make better decisions by thinking about future events so vividly as if we were experiencing the pleasure of the imagined rewards [4,5]. A decision to go to the gym might result from a mental simulation of rewarding experiences such as the thrill from getting yelled at by that energetic spinning instructor or the relaxing shower after the workout. Overwhelming evidence shows that both experienced and anticipated rewards are signalled by activity in the dopaminergic (DA) system [6], which also comprises the core brain reward circuitry including the ventral striatum and ventromedial prefrontal cortex (vmPFC [7]). These reward signals are thought to reflect learned associations between reward cues and reinforcers [6]. But what is remarkable from this wealth of research is the consensus that DA not only signals experienced but also purely anticipated rewards, confirming its pivotal role in decision making.

By contrast, very little is known about how simulation of physical effort could guide choice. Here two possible scenarios have been proposed. First, a prevailing idea from the effort discounting literature posits that efforts are represented as costs associated with the action [8]. These costs are thought to be compared with, or deducted from, the rewards to compute the subjective value of the action [8–10]. Similar to representation of experienced cost, any simulated cost signal would thus scale monotonically with increasing effort [11]. Second, choices may require simulation of the energization needed to ensure that the action can be successfully achieved [12]. Thus, simulated energization may draw on the same brain system that facilitates actual behavior energization [13]. The possible correspondence between anticipated and experienced effort-related energization in such brain signals may be analogous to how dopaminergic signals are equally sensitive to anticipated and experienced reward.

Teasing apart these two scenarios is not trivial. Both simulated cost and energization signals would scale monotonically with effort; however, one useful way may be to investigate how these signals differ between choice outcomes for the same effort/reward combination: “Yes” decisions in which individuals choose to engage versus “No” decisions in which individuals decide to forego the given effort. Any variation in choice outcome from trial to trial, given identical efforts and rewards at stake,

should reflect momentary fluctuations in the strength of reward and effort representations, allowing a closer inspection of whether stronger anticipatory effort signals indeed decrease the subjective value associated with option (for simulated cost signals) or signal higher readiness of the organism to take on this challenge (for simulated energization signals). In the former scenario, any neural effort signal should be higher in "No" compared to "Yes" decisions, consistent with the proposal that stronger representation of the effort-related cost decreases the value of the option and thus increases the chance of rejection. The second scenario, however, would predict the opposite pattern of results, with higher effort-related representations during "Yes" compared to "No" decisions. This is because a higher anticipatory energization signal would signal higher readiness of the organism to take on the challenge associated with the required effort level, thereby increasing the chance of acceptance of the choice option. While both of these influences of effort representations on choices are plausible, they make opposite predictions that we can test by comparing the strength of the corresponding neural signals during "Yes" and "No" decisions (see Fig. 1).

=== Figure 1 around here ===

Which brain systems may signal both experienced and simulated effort, just as DA and the core brain reward circuitry do with reward? The literature focuses mainly on noradrenergic arousal systems: While studies in rodents and monkeys show dominant DA encoding of upcoming rewards, hardly any systematic effects are documented for dopaminergic effort coding [14]. Moreover, neural signals related to rewards and effort appear to dissociate in terms of timing: Reward-linked firing of DA neurons increases during the decision process, whereas effort-linked noradrenergic (NA) activity is mostly observed *after* decision making, during the actual effortful action. In these situations, locus coeruleus (LC) neurons show activity increases that scale up with the size of effort that is currently being exerted [13]. These findings are usually taken as support that the NA system serves to optimize performance [15] by modulating arousal states [16,17] that provide neuromodulatory input to the entire neocortex [18,19]. Interestingly, such effort-linked NA activity can directly influence pupil dilation [20], making pupil width an accurate indicator of variability in multiple parameters for brain arousal states and behavioral performance [21]. Thus, the current literature mainly provides evidence that the experience of effort draws on pupil-linked NA arousal processes, which presumably mobilize the resources needed for behavioral energization [13]. Importantly, however, these data only pertain to experienced effort. It is therefore particularly interesting to test if the arousal system is also involved in simulating effort during the choice process, and whether the signal would play a role in simulated cost or simulated energization.

Which cortical areas may be affected by the arousal system during effort simulation? The answer to this question is unclear at present. Several human functional magnetic resonance imaging (fMRI) studies show anticipatory reward signals that are subjectively "discounted" by effort [8–10,22–25], consistent with the notion that the brain may encode physical effort as a type of cost [8–10,26]. However, this net

value signal may well reflect the rewarding aspects of the choice options, which blurs the interpretation whether this could reflect simulation of effort. By contrast, only few studies have identified signals for effort levels per se [22,23], some in SMA, ACC, and anterior insula for anticipated effort in non-choice settings [27,28], while others in the primary motor area and anterior insula for experienced effort [9,29]. Thus, while a neural representation of net value seems well established, there is little information about how the brain represents effort per se, either anticipated or experienced. The limited observations nevertheless suggest that activity in ACC/motor/insular network during the choice process may reflect effort simulation, which may be affected by arousal processes when simulating effort.

To shed light on all these issues, here we investigate systematically to what degree the arousal system may signal simulated effort during choice, and what behavioral function these signals may relate to. We first ask whether the arousal system, as indexed by pupil signals, encodes anticipated effort as a simulated cost or simulated energization. At the neural level, based on previous work on anticipated and experienced effort [9,27–29], we examine whether cortical representations of choice in the ACC/motor/insular network are modulated by effort amount and whether these effort-modulated choice representations link to arousal. To test for these effects, we measured phasic changes in pupil width during choice. Phasic pupil is a plausible candidate for signalling of effort simulation, since several studies show that pupil diameter increases during performance that requires mental [30] or physical effort [31]. One phasic pupil measure that has been particularly useful is the rate of pupillary dilation (ROD), which refers to the speed at which the pupil width changes within a certain period. Seminal work [20] in monkeys found the fastest rate of pupil dilation to occur 310ms ms after LC firing, suggesting a tight relationship between LC firing and not just pupil size but also the speed of dilation. In mice, both NA activity and cortical arousal states were more closely associated with rate of pupil change than with absolute pupil size [32]. Finally, in humans the rate of dilation was also associated with performance in a fast-paced sustained attention task [33]. We therefore focused on ROD as candidate marker of the speed with which arousal is upregulated and tested whether this reflects simulated cost or simulated energization.

In our study, we acquired pupil responses during fMRI of an effort/reward tradeoff choice task to identify anticipatory pupil and neural signals for efforts that have an impact on choice. First, we explored whether pupil-linked arousal during decision making, as measured in ROD, is associated with choices, in a manner that depends on the level of effort. In line with the two conflicting scenarios outlined in the literature, we reasoned that stronger responses for “No” decisions (reject effort) would support the view that effort is cognitively represented as a cost, whereas higher responses for a “Yes” decision (accept effort) would back the interpretation that effort representations signal behavioral energization for the future challenge. Second, we investigated whether we could find an analogous effort-modulated choice effect in neural activity, potentially within the ACC/motor/insular network, that could plausibly be

affected by noradrenergic arousal processes. Third, we tested for the correlation between these effort-modulated choice effects in the pupil data and brain activity. Fourth, if effort simulation is at all behaviorally relevant then we expect these pupil and brain responses to be associated with individual differences in how effort affects overall choice, as measured in an effort-discounting parameter derived by fitting a choice model to the behavioral data. Notably, we would expect the individual strength of effort discounting to correlate with different types of signals. In the simulated-cost scenario, we would expect effort discounting to be positively correlated with higher signal for "No" decisions (higher simulated cost), since individuals who assign higher costs to effort should reject the lotteries more often (high effort discounter). Under the simulated-energization scenario, however, we would expect effort discounting to be positively correlated with the higher energization signal for "Yes" decisions, since high effort discounters would need a stronger anticipatory energization signal to accept a given effort level.

Finally, to ascertain that our effort simulation effects were not driven by endogenous fluctuations of arousal states (rather than effort-linked trial-specific effects), we also examined tonic pupil as indexed pre-trial pupil baseline level (PBL). PBL is associated with choice variability [34], and elevated emotional arousal prior to a force-production task can also increase voluntary effort [35]. Thus, we conducted control analyses to test whether these pre-trial tonic arousal effects may cause a general bias towards exerting effort.

Results

Participants made decisions in the scanner about whether to *accept* or *reject* a reward offer (1 of 6 levels, from 0.50 to 10 CHF) that required exertion of physical effort (1 of 6 levels, from 40% to 90% maximum voluntary contraction--MVC) (Fig. 2). To ensure that participants would not treat the task as hypothetical decisions about trivial effort, we (1) devised a force task that mimics a typical strength exercise at the gym, with a cycle of 10 repetitions ('reps') of hand muscle contractions and relaxations for each effort level. As an illustration, we depict grip force traces from a training session (1 trial = 1 cycle of 5 'reps') done by one subject (Fig. 2C). We also (2) ensured that participants understood the real consequences of their decisions (they had to execute a random selection of eight choices after the scan). Rejecting the offer meant selecting a counteroffer of either 30 or 40% of the reward amount paired with the lowest force level (L-1). Critically, these decisions were temporally separated from the actual exertion (which happened after the experiment), to set up a hard test whether arousal effects could still be observed in cases where post-decisional motor preparation was completely absent. Given this experimental design, any phasic arousal effect could not be due to an impending motor action, and

any lack of such an effect would unlikely be due to the effort task being too trivial for the subjects. We could thus investigate whether pupil-linked arousal scales with increasing physical effort during mere mental simulation when deciding about future efforts.

Behavioral evidence for systematic effort-reward trade-offs

Initial analyses confirmed that participants indeed systematically traded off the proposed efforts and rewards when taking choices, as could be expected based on previous work [9,10,36]. Offers were accepted significantly more often when they were coupled with higher rewards (logistic regression of choice (accept =1/reject=0), n=49; $t_{reward}(48)$=6.93, p<0.0001) and lower effort ($t_{effort}(48)$=-7.25, p<0.0001). Offers were accepted / rejected particularly often when they were clearly attractive (high rewards for low effort) / unattractive (low rewards for high efforts) ($t_{reward*effort}(48)$=-1.93, p=0.06; Fig. 2D). These choice effects were also corroborated by the response time (RT) data. Clearly bad (low reward, high effort) and clearly good offers (high reward, low effort) were associated with faster responses (Fig. 2E). More specifically, RTs were not only influenced significantly by the offered levels of reward (multiple regression of RT (z-scored), n=49: $t_{reward}(48)$=3.93, p=0.0003) and effort ($t_{effort}(48)$=-5.90, p<0.0001), but were also faster when participants accepted than when they rejected the offers (t_{choice} (48)=-4.46, p<0.0001; rightmost plot in Fig. 2E; other effects: $t_{choice*reward}(48)$=-5.82, p<0.0001; $t_{choice*effort}(48)$=8.44, p<0.0001; $t_{constant}(48)$=6.68, p<0.0001; $t_{reward*effort}(48)$=-0.8, p=0.41; $t_{choice*reward*effort}(48)$=1.3, p=0.019). These results confirm previous findings that decisions vary as a function of the offered rewards, the required effort, and the decision outcome.

=== Figure 2 around here ===

Rate of pupillary dilation during choice reflects simulated energization

Pupillary responses during decision making showed a stereotypical dilation shortly following cue onset, peaking right after response onset, and constricting down to baseline level around cue offset (Fig. 3A). To examine whether anticipated effort indeed engages the arousal system during choice, we compared the rate of pupil dilation (ROD) for trials in which participant accepted vs rejected offers (“Yes” vs “No” decisions, respectively) that required low, middle, or high effort (3x2 effort-by-choice repeated measures ANOVA, n=42). This revealed that the pupil dilated significantly faster when participants accepted (versus rejected) an offer comprising a high effort ($F_{effort\text{-}by\text{-}choice}(2,82)$=3.81, p=0.02). This effect was specific to high-effort trials (comparison of accept versus reject for high-effort trials: $t_{high\text{-}effort}(41)$=2.39, p=0.02; $t_{mid\text{-}effort}(41)$=1.40, p=0.90, $t_{low\text{-}effort}(41)$=0.12, p=0.90; Fig. 3B). Thus, the effort-

modulated choice effect in the pupil signal is consistent with the scenario that arousal system engagement during choices about future efforts relates to behavioral energization for a challenge in the future. Importantly, a comparable mirrored effect for low-effort trials (reject > accept low effort) was not significant (p=0.90). This shows that the pupil-dilation effect for accepting high-effort trials cannot reflect errors, infrequent occurrences, or surprise [37], which would be similarly present for accepting high-effort and rejecting low-effort options.

To investigate further the specificity of these links between choices to accept high effort-options and ROD, we controlled for all other variables in our design within a logistic regression of choice (accept =1/reject=0, n=49). This replicated the effects of reward (t(48)=6.61, p<0.0001), effort (t(48)=-7.39, p<0.0001), and their interaction (t(48)=-2.43, p=0.01; $t_{constant}$(48)=4.21, p=0.0001) but crucially also revealed a significant ROD-by-effort interaction (t(48)=2.23, p=0.03), all other effects are ns (ps>0.05; fig. 3D). Thus, the simulated energization signal visible in the pupil dilation cannot be accounted for by other variables in our experimental design. Please note that in this regression, we also included pupil baseline level (PBL) and its interaction with reward and effort; endogenous arousal fluctuations prior to stimulus onset were thus controlled for and could not bias our results.

While this extended regression model highlights a novel association between ROD, effort, and choice, it may well be that this pupil measure does not add significant predictive information on top of what can be extracted from the reward and effort associated with the present choice option. To test this, we compared a measure of model-fit (adjusted R-squared) between the extended 'ROD' regression model and a classical 'null' regression model with only reward, effort, and the interaction, but no pupil measure. As seen in Fig. 3D, there is a higher model fit for the extended 'ROD' regression compared to that for the classical 'null' model. This result suggests that ROD explains additional variance in the choice data, above and beyond what can be accounted for only by reward, effort, and the interaction (n=49; t(48)=4.80, p<0.0001).

Simulated energization in pupil relates specifically to effort-reward trade-offs

To investigate whether the simulated energization process during choice is indeed behaviorally relevant (i.e., systematically linked to effort-reward trade-offs), we tested whether the ROD energization effects were associated with individual differences in effort discounting. For this analysis, we employed each individual's effort discounting parameter from a parabolic effort discounting model (selected as the best model from 8 competing models based on random-effects Bayesian model comparison [38] see supplementary Fig. S4). This subject-wise parabolic effort discounting parameter was indeed significantly correlated with the effect in ROD (n=42; r(40)=0.34, p=0.02; robust regression; b(40)=0.13, p=0.03; fig. 3C). Thus, subjects with higher effort discounting (i.e., whose overall choice was more

strongly affected by increasing effort) indeed showed faster pupil dilation when accepting compared to rejecting high effort. These results fit well with the finding that the cost of effort is represented in a non-linearly increasing manner as the effort amount increases, captured by a parabolic discounting shape [10,39,40]. This non-linearity is also evident in our observation that the energization effect was only evident for high-effort trials, comprising the most difficult effort levels (80-90% of maximum force). Importantly, across subjects, we find that the energization responses in both pupil and the brain are positively associated with a subject-specific parabolic effort discounting parameter, consistent with the idea that this signal may be relevant for guiding overall choice.

Simulated energization effects in pupil are independent of decision difficulty

Despite the tight relationship between the energization signals evident in the pupil and effort discounting, it is theoretically possible that the energization effect we observe in pupil signals may not just relate to choice outcome but may also be higher for trials that are subjectively difficult, as larger pupil size has been observed for trials that require greater cognitive control [30]. This effect might be confounded with the energization effect, particularly because in some cases, high effort trials may be associated with high rewards, hence making the decision to either select or forego effort more difficult. To investigate this possibility, we directly quantified decision difficulty by calculating trial-wise absolute difference in subjective values between the two options on each trial (dSV; see Methods), with smaller values indicating harder decisions. In addition, we also inspected RTs, which are commonly used as indirect proxy for decision difficulty [41]. Indeed, we found significant choice-by-effort interaction effects on both proxies of decision difficulty (dSV and RT; 3x2 repeated measures ANOVA), suggesting that the ROD effects we report may share variance with direct and indirect proxies for decision difficulty. Therefore, to directly investigate whether this simulated energization effect is clearly independent of choice difficulty, we repeated the analyses reported above (and depicted in Fig 3B-E) on the residuals of ROD after partialing out the effects of dSV and of RT (orthogonalization of ROD relative to these variables, one at a time). Encouragingly, these control analyses revealed the very same effects already shown in Fig 3, namely (1) significantly higher residual ROD in accept versus reject high-effort condition, (2) significant effort-by-residual ROD effect on choice, (3) significantly higher model fits for the extended regression models with residual ROD compared to a classical 'null' model, and (4) significantly positive correlations between the size of the energization effect of the residual ROD (accept minus reject high-effort) with the parabolic effort discounting parameter (Fig. S5). Thus, the simulated energization effect we identified in ROD is independent of decision difficulty and reflects different neural mechanisms to those underlying conflict-driven pupil dilations and behavioral adjustments [42].

=== Figure 3 around here ===

Neural evidence for systematic effort-reward trade-offs

Concurrent with behavioral evidence that participants systematically trade off reward with effort, we examined the neural representations of reward, effort, and the interaction. In addition, we also examined brain activity correlating with ROD. We replicated previous findings [7,25] of neural reward representation in the ventral striatum and effort modulation in the frontal pole (FWE p<0.05; Fig S6). In addition, using another GLM, we replicated previous finding that brain activity in the vmPFC is correlated with the computed subjective value based on the amount of reward that is subtracted by the amount of effort (FWE p<0.05; Fig S7). Taken together, our brain results fully replicate previous data identifying cortical and subcortical brain regions that support effort-reward trade-offs.

Arousal-linked simulated energization is reflected in neural responses in SMA/ACC and anterior insula

We next examined how the behavioral energization identified in the pupil signals relates to modulations of neural effort representations during the decision process, by running a two-way within-subject ANOVA with effort (low, mid, high bins) and choice (accept, reject) of the brain responses to the presentation of the options (cue onset). We specifically tested for the neural version of the simulated energization signal we observed in the pupil data, i.e., a significant activity increase specific to the decision to accept high-effort trials. Such positive effort-by-choice interactions for BOLD activity were revealed in right anterior insula, left anterior insula, left ACC (extending to the SMA) (MNI space coordinates: [33, 24, 3], [-33, 24, 0], [-9, 24, 33]; peak *F* values, 22.10, 14.77, 20.75; extent: 127, 71, 350 voxels; $p< 0.0001$, $p=0.002$, $p< 0.0001$ FWE, respectively; Fig. 4A; GLM1a in Methods), along with activations in bilateral caudate and midbrain (full statistics and results for the main effects are found in Table 1). To assess the specificity of this effect for high-effort trials, we tested for simple effects of choice for all different effort levels. This confirmed higher activity in the same ACC-anterior insula clusters, along with activity in nucleus accumbens, only when participants accepted versus rejected high-effort offers (FWE p<0.05; Table 1), but not for the other effort levels.

To link the neural simulated energization responses with the corresponding signal in pupil, we correlated the amplitudes of the neural response to accept>reject high-effort trials with the same accept>reject high-effort contrast in ROD ('ROD energization'; Fig. 4B). This revealed the same circuitry of ACC and bilateral anterior insula observed for the behavioral effect only (see above), corroborating that those participants who had faster pupil dilations when accepting (vs rejecting) high effort also had higher choice-related brain activity in these regions (Fig. 4C; Table 2). This positive relationship was

also confirmed in an analogous ROI analysis, correlating the simulated energization pupil and neural measures extracted from functional ROIs of ACC and bilateral anterior insula that were independently defined from the choice-by-effort F contrast ($r_{ACC}(40)=0.58$; $r_{L.Insula}(40)=0.52$, $r_{R.Insula}(40)=0.64$; ps<0.0001; Fig. 4D).

Finally, given that the energization effect in the pupil dilation data was correlated with the individual effort-discounting parameter (Fig 3E), we also inspected whether the brain responses for the decision to accept high effort would be associated with individual differences in effort discounting. To test this, we again took each individual's parabolic effort discounting parameter and used it as a subject-specific covariate at the second level for our critical contrast (accept>reject in high effort bin). This confirmed that the neural measure for simulated energization was correlated with participants' effort discounting (Fig. 4E-F; Table 2) in ACC and within the same functional ROIs defined above ($r_{ACC}(40)=0.61$, p<0.0001; $r_{L.Insula}(40)=0.35$, ps=0.02; $r_{R.Insula}(40)=0.48$; p=0.0011). Thus, like the pupil-related arousal signals, neural responses in these areas during choices to accept high-effort trials were strongest in people with higher effort discounting.

Taken together, our data show that brain activity in ACC and anterior insula shows anticipatory effort signaling in a way that is consistent with simulated energization for high physical challenges. These areas show higher activity during decisions to take on a difficult physical task in the future, and this activation is tightly linked to anticipatory activation of the arousal system and to the weight that participants place on effort when trading off rewards and efforts during choice.

=== Figure 4 around here ===

Finally, to ascertain that decisions were not driven by the ongoing level of background arousal, we defined the average pupil diameter during 500 ms prior to the presentation of the options, as an index of pre-trial pupil baseline level (PBL). We contrasted choices for which participant accepted or rejected offers that required low, middle, or high effort (tertile split; 3x2 repeated measures ANOVA; n=42). We found no significant difference in PBL between choices to accept or reject ($F(1,41)=0.16$, p=0.69) and no effect of the different effort levels (main effect: $F(2,82)=0.76$, p=0.47; effort-by-choice interaction: $F(2,82)=1.37$, p=0.26; Fig S3A). This absence of a link between PBL and effort-based choice did not reflect more complex interactions with other experimental factors or influences from the previous trial, as ascertained by logistic regressions of choice on PBL, RT, reward, effort, and the interactions (no significant effect, see Fig S2B-C). Thus, we found no evidence that ongoing background arousal state, as indexed by pre-trial pupil baseline, would bias subjects to accept high-effort options, thus confirming the specificity of the energization effect for phasic arousal responses during the choice process.

Discussion

We examined how the brain may represent future efforts during choice, motivated by the wealth of data on how it represents effort level during actual physical exertion. Specifically, we directly tested two competing hypotheses against one another: Whether such neurobiological representation of future effort signals simulated cost or energization. Consistent with the latter, our results show stronger activity in the arousal system (as measured in pupil) and ACC-insular brain network for choices that involve anticipating a sizeable amount of effort. This emphasizes that future effort during choice is represented by arousal system in a way that appears to relate to future energization.

Our results emphasize that phasic pupil-linked arousal during the decision process is tightly linked to choice outcome, but they also raise the question what neural mechanisms may lie at the heart of this link between behavior and neural signals. There are at least two plausible answers to this question. First, simulating the required energization could have a “bottom-up” influence on decisions to produce a bias towards accepting effort. This would be consistent with the widely held view [43] that the strength of neural representations for choice attributes directly influence the decision – for instance, it has been shown that intensifying encoded rewards through simulation of future episodic events is linked with decisions that promote higher long-term pay-offs [4,5] and even increases prosocial behavior [44]. Given this assumption, the arousal signal we observed in this study might either down-modulate effort encoding or shift the decision rule [45], implying that a sufficiently strong arousal signal could bias a decision towards taking on the physical challenge. As for neural implementation, phasic LC activity is known to transmit feedforward information to ACC via ascending projections to prefrontal (PFC) [18,19,46], providing a plausible pathway for such bottom-up influences. Nervous readout of the autonomous activation associated with arousal could provide an additional mechanism by which the arousal signal observed here may bias choices, serving as a signal that the organism is indeed ready to take on the physical challenge.

Second, simulated energization could simply be a byproduct of choice, implying a top-down influence from the cortical decision circuit to the arousal system. Decision outcomes could be relayed in the form of cortical descending input from the PFC into LC. ACC activity has been coupled with pupil diameter [42,47] and the timing of pupil modulation by ACC in some cases precedes that by LC [20]. Existing tracing data in rodents and monkeys also show afferent PFC projections as the main direct cortical influence on LC [48,49]. Intracranial stimulation in human ACC leads to subjective accounts of changes in arousal states, such as increased heart rate, coupled with the anticipation of challenges and a strong motivation to overcome it [50]. This interpretation is also closely linked, though not identical, with the proposal that ACC computes the expected value of mobilising mental resources [51]. Taken together,

these observations are consistent with the idea of a top-down influence from ACC to NA arousal system [52], which may serve to transmit information about the commitment to overcome great physical demand, thus resulting in automatic speeded upregulation of arousal states to prepare the organism for the future challenge associated with the recent choice.

Although our current study cannot give a conclusive answer on which of these two alternative explanation holds, in our data arousal does not seem to exert any bottom-up modulation of neural effort representations that could allow arousal to instantaneously bias valuation. In addition, we did not find evidence that baseline fluctuations of arousal prior to the presentation of the options played any role in decisions. Instead, the phasic arousal signals we observe seem to relate systematically to activity within the cortical decision circuit, consistent with the notion that the brain simulates the already-selected effort by means of arousal signalling. However, future studies may need to employ neuroimaging methods with higher temporal resolution to disambiguate fully these two hypotheses. Such studies may also employ pharmacological manipulation to increase NA tone activity, bio/neuro-feedback with pupil/LC activity, and mental simulation training [53] to increase arousal in a bottom-up fashion.

What would be the cognitive purpose of simulating behavior energization associated with a choice? Such simulation may contribute to metacognitive processes that evaluate the quality of our ongoing decisions to optimize future decision making [54]. For an example from another domain, there is evidence that actual experience of choice and success in obtaining a food item influences how we value the food item in the future [55]. Effort simulation may thus serve as a rich milieu for 'scene construction' [56] in which subjects evaluate the quality of their decision, which has the potential to shift future valuation. In our context, the source of simulation may include drawing from memory how much cognitive control needs to be mobilized [51] in order to keep exerting physical effort rather than quitting, or retrieving the memory of previously incurred metabolic signal that accumulated the longer subjects exerted physical effort [29]. Future experiments may directly test this conjecture by devising mental simulation paradigms in which participants imagine these specific elements of the force task, namely the sensations of mental fatigue or pain, and assessing how vividness ratings of these imagined bodily sensations would correlate with brain activity and choice. Furthermore, a mental simulation paradigm that manipulates agency might reveal stronger simulation signals for one's own decisions compared to experimenter-imposed decisions, which would lend evidence for the use of simulation for self-evaluation [54].

Irrespective of these considerations, our results highlight a plausible partnership of the dopaminergic and noradrenergic systems in anticipatory reward and effort processing guiding choice. The majority of effort studies so far (including our current data—see FigS6-7) have reported a net value representation (reward discounted by effort) within the core brain valuation network [9,22], and in dorsal PFC areas including SMA/ACC [8–10,23–25]. These fMRI results are consistent with animal data showing reduced willingness to choose a high-effort/high-reward option when dopamine is depleted [57] and with

the overarching dopaminergic role in upcoming and ongoing motivational reward processing [14]. Here our data support the intriguing view that upcoming effort may be represented by the same brain and arousal mechanisms previously linked with ongoing physical effort, involving SMA/ACC and anterior insula and NA-originated pupil dilations [13,27–29,31]. This partnership, DA for reward and NA for effort, does not seem to correspond with the classical but possibly simplistic view that DA-linked reward processing is discounted in a subtractive fashion by NA-linked effort cost representations. However, we emphasize that our behavioral data and some aspects of our neural results clearly concur with previous findings that an option is selected based on a trade-off between reward and effort (FigS7). What has been unexplored in previous fMRI work, however, is how the noradrenergic arousal system is sensitive to effort, and in what way this neurobiological representation of effort is functional for choice. Using concurrent pupil-fMRI in an effort discounting task, we were able to scrutinize the precise functional role of NA in signalling future effort in humans, and indeed, our results suggest that NA seems to show a complementary function to DA, potentially allowing the organism to follow through DA-driven decision arbitrage processes by means of arousal signaling that ensures appropriate NA-driven behavioral energization in the future [52,58].

Variations in arousal states (measurable by pupil activity) - such as locomotion and sleeping - are coupled with oscillatory state changes in brain networks [17] and these are thought to result from noradrenergic innervation to the cortex [46]. However, there are also observations that cholinergic neuromodulatory projections from the basal forebrain to the cortex are intimately associated with movement during wakefulness and REM sleep [59], which is often confounded with arousal states. This raises the concern whether we can truly draw the conclusions that our arousal effects evident in the pupil signals originate from NA-LC neuromodulation. While we cannot fully rule out the effects of cholinergic activity, a recent analysis with pupil activity and noradrenergic and cholinergic projections shed light on this issue, demonstrating that rate of pupillary dilation in mice is more tightly linked with NA projections to the cortex, whereas activity in the cholinergic pathways more closely matched absolute pupil diameter [32]. These data support the view that our ROD effects reflect phasic arousal variations that most likely originated from NA-LC activity.

Our results may have relevance for the diagnosis and therapy of brain disorders with deficits in motivated behavior. Committing to effort is a first step for success in motivated behaviors and the inability to commit to effort may bring about a cascade of clinical symptoms of apathy with a core feature of lack of self-initiated actions [60]. Recent neurocomputational work on effort-reward tradeoffs has identified promising phenotyping approaches of motivation disorders; these reflect key involvement of the fronto-subcortical circuitry and neuromodulatory systems including dopamine, serotonin, and noradrenaline [61,62]. A specific role for noradrenaline is suggested by the finding that motivation deficits in depression that are inadequately treated by serotonergic antidepressants – including fatigue and

loss of energy – have been shown to significantly improve following administration of NA (and dopaminergic) agents [63]. This highlights the critical yet overlooked role of NA in motivation regulation in depression [64]. Our study contributes to this body of work showing that the pupil-brain arousal system is sensitive to deliberations regarding sizable intensities of physical effort. Future work may focus on further incorporation of autonomic arousal and noradrenergic systems in quantitative models of motivation deficits [62], particularly in dissociating arousal effects of effort from the more commonly known effects of reward.

Materials and Methods

Participants

Fifty-two right-handed participants (29 females, mean age=22.3 (3) years) volunteered to participate in this study. Participants were informed about all aspects of the experiment and gave written informed consent. They received between 80-100 CHF (depending on the realized choices and performance) for their participation. Participants were screened for MRI compatibility. They had no neurological or psychiatric disorders and needed no visual correction. The experiments conformed to the Declaration of Helsinki and the protocol was approved by the Ethics Committee of the Canton of Zurich. Data from one subject were excluded because of eye tracker data loss. Inclusion of this subject in the behavioral analysis did not change the statistical results, but for consistency we excluded this data set from all analyses. We then screened subjects based on their mean choice proportion to be within 0.1 and 0.9, thus excluded data from one subject whose rate of acceptance was 0.95. The final N was 49. However, in certain analyses in which we had to split the data in accordance with our critical pupil contrast, we had 7 subjects with certain data bins missing. Given the specific emphasis on the effects seen in pupil, we were therefore only able to conduct the neuroimaging analysis with n=42.

Procedure

Upon arrival, participants were seated in the behavioral testing room, filled the MRI screening and consent forms, and received general instructions on the force task and MRI safety. Maximum voluntary force (MVC) level for each hand was obtained by averaging the top 33% force values produced during three 3-s squeezes. Continuous vocal encouragement was given during entire squeeze period (e.g., “keep going, keep it up”).

Guided by a vertical bar on-screen (Fig. 1A), participants were trained to do set squeezes from force levels of 10%-90% MVC (shown to subjects as level 1-9), alternating between left and right hand. One set consisted of 5 repetitions (‘reps’) that lasted 3 s interleaved by 3 s rest periods. Participants

experienced all levels from 1-8 once, randomly assigned to either left and right, and level 9 twice, once for each hand. The order of force levels was pseudo-randomised. Half of the subjects practiced on levels 1, 3, 5, 7, 9 with left hand and 2, 4, 6, 8, 9 with right hand, and vice versa for the other half of subjects.

Following a 5-minute break, they proceeded with a subjective rating task in which they had to squeeze for each hand once at levels 1, 3, 5, and 9 for 5 s without knowing the difficulty levels and rated on a continuous visual analogue scale how effortful the grip was for them. They were explicitly instructed that the leftmost and rightmost point in the scale should refer to level 0 and level 10, respectively. Mean pearson's *r* between subjective ratings and the object force levels were 0.93 (sem=0.0073), one-sample *t*-test against *r* of 0: $t(46)=127.63$, $p<0.0001$, suggesting a close relationship between subjective and objective effort and successful force training.

Prior to scanning, participants made five practice decisions and we made sure that participants fully comprehended the task. They were also fully aware that 8 randomly selected decisions (of 10 'reps' each time, rather than the practiced 5 'reps'), would be implemented in the behavioral testing room after the scan.

Effort Discounting Task

Participants made decisions between performing a specific effort level of the force task (between levels 4-9) to earn varying reward amounts (0.5, 1, 3, 5, 8, 10 CHF) and performing a counteroffer force task at level 1 to earn either 30% or 40% of the reward of the first offer (Fig. 2C). The force task involves performing one set of 10 'reps' at the selected effort level. Participants were fully aware that they would make successive decisions in the scanner without executing the force task and they were not provided with the dynamometer.

We used a factorial design, with six effort and six reward levels (36 cells), and two reward counteroffers per cell (3 exemplars each), totalling in 216 trials. Trials were split in three fMRI runs of 72 trials (9 mins); trial order was pseudorandomised per subject per run.

During a fixation period of 3-6 s (created using the function gamrnd(0.8,1), mean 3.7s), the text indicating reward and effort levels were masked with a series of letters "X" (Fig. 1B). Following this period, the colour of the + sign at the centre changed and the effort and reward of each of the two options were presented on either side of the fixation point for a fixed duration of 3 s. This prompted the subjects that they were able to press either the left or the right key to indicate their choice. To provide decision feedback, key response was promptly followed by a change in colour for the selected option.

Pupillometry

Participants' right or left eye (depending on feasibility) was monitored using MR-compatible infrared EYElink 1000 eye-tracker system (SR Research Ltd.) with 500 Hz sampling rate. Participants were instructed not to blink during the presentation of the options. Pre-processing of the pupil data was

performed in Matlab (version 2017a, MathWorks, Natick, USA). Data indicating eye blinks were replaced using linear interpolation. The data were visually inspected to ensure that all artefacts had been successfully removed. Pupil data were z-transformed within each run to control for variability across runs and across subjects. Rate of dilation (ROD, unit: std/s), one of our measures of arousal, was calculated by subtracting pupil size at button response from pupil size at cue onset, divided by response times. Pre-trial pupil baseline level (PBL) was calculated by averaging pupil size from 500ms - 1ms before stimulus onset.

To ensure constant screen luminance level, we kept roughly the same number of pixels throughout the events by replacing the text indicating reward and effort levels with a series of Xs and by using text hues that were isoluminant to the grey background (RGB grey: 178.5, 178.5, 178.5; green: 50, 100, 10; purple: 118, 60, 206; blue: 53 77 229). Ensuring readability, we selected these hues out of 17 theoretically isoluminant hues where relative luminance was calculated as a linear combination of the red, green, and blue components based on the formula: Y = 0.2126 R + 0.7152 G + 0.0722 B. This formula follows the function that green light contributes the most to perceived intensity while blue contributes the least (Stokes, et al.; https://www.w3.org/Graphics/Color/sRGB). Green was always fixed as the base hue and blue and purple were randomly assigned trial-by-trial to highlight the selected offer (Fig. 1B).

Additionally, in a control experiment, we recorded luminance-driven pupil dilation without any cognitive task. We presented fixation screens with a series of Xs as fixation period and Ys to replace the text that would have indicated the effort and reward levels in the main experiment, each period lasting for 3 s. Participants were instructed to keep their eyes open but were not required to press any key. Just like in the main experiment, green was the base hue during fixation whereas blue and purple were used to highlight the text on one side of the screen. All stimuli were in the same text format as in the main task (Fig. 2B). Order of hue and side assignment were all counterbalanced and pseudorandomised. We found no difference in mean pupil diameter during the presentation of these control stimuli in different hues, confirming that the pupil response in the main task was not driven by differences in text luminance (Fig. S1).

fMRI Acquisition and Analysis

Functional imaging was performed on a Philips Achieva 3T whole-body MR scanner equipped with a 32-channel MR head coil. Each experimental run contained 225-244 volumes (voxel size, 3x3x3 mm^3; 0.5 mm gap; matrix size, 80x78 (FoV: [240 140 (FH) 240]; TR/TE 2334/30 ms; flip angle, 90°; parallel imaging factor, 1.5; 40 slices acquired in ascending order for full coverage of the brain). We also acquired T1-weighted multislice gradient-echo B0 scans which were used for correction of deformations (voxel size, 3 x 3 x3 mm^3; 0.75 mm gap; matrix size, 80x80; TR/TE1/TE2 // 400/4.3/7.4 ms; flip angle, 44°; parallel imaging; 40 slices). Additionally, we acquired a high-resolution T1- weighted

3D fast-field echo structural scan used for image registration during postprocessing (170 sagittal slices; matrix size, 256x256; voxel size, 1x1x1 mm3; TR/TE/TI // 8.3/3.9/1098 ms).

We used Statistical Parametric Mapping (SPM12; Wellcome Trust Centre for Neuroimaging, London, http://www.fil.ion.ucl.ac.uk/spm) for imaging analyses. Four preprocessing steps included realignment and unwarping, slice-timing correction, coregistration and normalization, and smoothing, and correction for physiological noise, these are described in supplementary materials.

We performed random-effect, event- related statistical analyses. For each subject, we first computed a statistical general linear model (GLM) by convolving series of stick functions, time-locked to the cue onsets, with the canonical hemodynamic response functions and their first derivatives (temporal derivative). We also added to these GLMs 18 physiological regressors and 6 motion parameters. At the second level, we then tested the significance of subject-specific effects (as tested by t-contrasts at the first level) across the population. For these analyses, we used a grey matter mask as an explicit mask, created by averaging across subjects and smoothing (8mm) all participants' normalized grey matter images (wc1*.nii) from the 'segment' procedure.

We built three first level GLMs. In GLM1, to highlight activity correlating with the interaction between choice (accept vs reject) and effort levels (low, mid, high bins), we defined six first-level regressors of interest representing the six different event types at cue onset: reject low effort (L0), accept low effort (L1), reject mid effort (M0), accept mid effort (M1), reject high effort (H0), and accept high effort (H1). To account for effects of RT, ROD, and reward, these varying indices were entered as trial-wise parametric modulators (z-scored) for each regressor. From this first-level GLM, we created 3 second-level GLMs focusing only on evoked responses at cue onset. In GLM1a, we entered the contrast images of all six regressors (against baseline) into a second-level 3x2 (effort bin x choice) within-subject ANOVA in SPM. We created GLM1b to inspect the association between neural and pupil effects, by entering the 'neural energization' (H1>H0) contrast images into second level one-sample t-test as a second-level subject covariate 'ROD energization' (H1 minus H0 in ROD). In GLM1c we used the same H1>H0 contrast and entered as subject covariate the effort-discounting parameter from computational modelling. To identify unique variance associated with each of our trial parameters, we generated GLM2 without any orthogonalization. We used the cue onset as a single regressor with choice (1=accept; 0=reject), z-scored reward, effort, reward-by-effort, ROD-by-reward, ROD-by-effort, and RT as trial-wise parametric modulators. Finally, to specifically replicate previous results on the neural representation of subjective value (SV), we built GLM3. We used the cue onset as a single boxcar regressor with RT as duration and z-scored SV of the offer as the only trial-wise parametric modulator. We computed SV using the reward and effort amounts of the offer of each trial and subject-wise discounting parameter from the winning model (parabolic effort discounting; FigS4). For both

585 GLMs 2-3, we then entered the contrast images of each parametric modulator vs baseline into second
586 level one-sample t-tests.

587 ***Statistical Analysis***

588 Statistical analyses for behavioral and pupil data were done with MATLAB 2012
589 (www.mathworks.com). We conducted (multiple) logistic or linear regressions separately for each
590 participant and entered the regression weights of each predictor from all participants into a one-sample
591 t-test. All continuous predictors were z-scored across trials within each participant. This approach
592 allows for the intercept (constant) to vary across participants. We ran two-way repeated measures
593 ANOVAs, with significant interactions followed up by paired-samples t-tests to examine simple effects
594 of one variable at each level of the other variable. We also used Pearson's correlations to test the
595 association between our critical contrasts with possible covariates. Computational modeling and further
596 statistical tests are describe in supplementary materials.

597

References and Notes

1. Bautista, L. M., Tinbergen, J. & Kacelnik, A. To walk or to fly? How birds choose among foraging modes. *Proc. Natl. Acad. Sci. U.SA.* **98**, 1089–1094 (2001).
2. Suddendorf, T., Addis, D. R. & Corballis, M. C. Mental time travel and the shaping of the human mind. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 1317–1324 (2009).
3. Bulganin, L. & Wittmann, B. C. Reward and novelty enhance imagination of future events in a motivational-episodic network. *PLoS One* **10**, 1–18 (2015).
4. Benoit, R. G., Gilbert, S. J. & Burgess, P. W. A neural mechanism mediating the impact of episodic prospection on farsighted decisions. *J. Neurosci.* **31**, 6771–6779 (2011).
5. Peters, J. & Büchel, C. Episodic Future Thinking Reduces Reward Delay Discounting through an Enhancement of Prefrontal-Mediotemporal Interactions. *Neuron* **66**, 138–148 (2010).
6. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science (80-.).* **275**, 1593–9 (1997).
7. Bartra, O., McGuire, J. T. & Kable, J. W. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* **76**, 412–27 (2013).
8. Klein-Flügge, M. C., Kennerley, S. W., Friston, K. & Bestmann, S. Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off. *J. Neurosci.* **36**, 10002–10015 (2016).
9. Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M.-L. & Dreher, J.-C. Separate valuation subsystems for delay and effort decision costs. *J. Neurosci.* **30**, 14080–14090 (2010).
10. Chong, T. T.-J. *et al.* Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLOS Biol.* **15**, e1002598 (2017).
11. McGuire, J. T. & Botvinick, M. M. Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7922–6 (2010).
12. Paravlic, A. H. *et al.* Effects and Dose–Response Relationships of Motor Imagery Practice on Strength Development in Healthy Adult Populations: a Systematic Review and Meta-analysis. *Sport. Med.* **48**, 1165–1187 (2018).
13. Varazzani, C., San-Galli, a., Gilardeau, S. & Bouret, S. Noradrenaline and Dopamine Neurons in the Reward/Effort Trade-Off: A Direct Electrophysiological Comparison in Behaving Monkeys. *J. Neurosci.* **35**, 7866–7877 (2015).
14. Walton, M. E. & Bouret, S. What Is the Relationship between Dopamine and Effort? *Trends Neurosci.* **42**, 79–91 (2019).
15. Yerkes, R. M. & Dodson, J. D. The relation of strength of stimulus to rapidity of habit-formation in the kitten. *J. Comp. Neurol. Psychol.* **18**, 459–482 (1908).
16. Pfaff, D. W., Martin, E. M. & Faber, D. Origins of arousal: Roles for medullary reticular neurons. *Trends Neurosci.* **35**, 468–476 (2012).

17. Takahashi, K., Kayama, Y., Lin, J. S. & Sakai, K. Locus coeruleus neuronal activity during the sleep-waking cycle in mice. *Neuroscience* **169**, 1115–1126 (2010).
18. Porrino, L. J. & Goldman-Rakic, P. S. Brainstem innervation of prefrontal and anterior cingulate cortex in the rhesus monkey revealed by retrograde transport of HRP. *J. Comp. Neurol.* **205**, 63–76 (1982).
19. Chandler, D. J., Lamperski, C. S. & Waterhouse, B. D. Identification and distribution of projections from monoaminergic and cholinergic nuclei to functionally differentiated subregions of prefrontal cortex. *Brain Res.* **1522**, 38–58 (2013).
20. Joshi, S., Li, Y., Kalwani, R. M. & Gold, J. I. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron* **89**, 221–234 (2016).
21. Yüzgeç, Ö., Prsa, M., Zimmermann, R. & Huber, D. Pupil Size Coupling to Cortical States Protects the Stability of Deep Sleep via Parasympathetic Modulation. *Curr. Biol.* (2018). doi:10.1016/j.cub.2017.12.049
22. Aridan, N., Malecek, N. J., Poldrack, R. A. & Schonberg, T. Neural correlates of effort-based valuation with prospective choices. *Neuroimage* **185**, 446–454 (2019).
23. Arulpragasam, A. R., Cooper, J. A., Nuutinen, M. R. & Treadway, M. T. Corticoinsular circuits encode subjective value expectation and violation for effortful goal-directed behavior. *Proc. Natl. Acad. Sci.* **115**, E5233–E5242 (2018).
24. Bernacer, J. *et al.* Neural correlates of effort-based behavioral inconsistency. *Cortex* **113**, 96–110 (2019).
25. Burke, C. J., Brünger, C., Kahnt, T., Park, S. Q. & Tobler, P. N. Neural Integration of Risk and Effort Costs by the Frontal Pole: Only upon Request. *J. Neurosci.* **33**, 1706–13 (2013).
26. Hull, C. L. *Principles of Behavior: An Introduction to Behavior Theory*. (Appleton-Century-Crofts, 1943).
27. Kurniawan, I. T., Guitart-Masip, M., Dayan, P. & Dolan, R. J. Effort and Valuation in the Brain: The Effects of Anticipation and Execution. *J. Neurosci.* **33**, 6160–6169 (2013).
28. Skvortsova, V., Palminteri, S. & Pessiglione, M. Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates. *J. Neurosci.* **34**, 15621–15630 (2014).
29. Meyniel, F., Sergent, C., Rigoux, L., Daunizeau, J. & Pessiglione, M. Neurocomputational account of how the human brain decides when to have a break. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2641–2646 (2013).
30. van der Wel, P. & van Steenbergen, H. Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychon. Bull. Rev.* **25**, 2005–2015 (2018).
31. Zénon, A., Sidibé, M. & Olivier, E. Pupil size variations correlate with physical effort perception. *Front. Behav. Neurosci.* **8**, 1–8 (2014).
32. Reimer, J. *et al.* Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat. Commun.* **7**, 13289 (2016).
33. Van Den Brink, R. L., Murphy, P. R. & Nieuwenhuis, S. Pupil diameter tracks lapses of attention. *PLoS One* **11**, 1–16 (2016).
34. Murphy, P. R., Vandekerckhove, J. & Nieuwenhuis, S. Pupil-Linked Arousal Determines Variability in

Perceptual Decision Making. *PLoS Comput. Biol.* **10**, (2014).

35. Schmidt, L. *et al.* Get aroused and be stronger: emotional facilitation of physical effort in the human brain. *J. Neurosci.* **29**, 9450–7 (2009).
36. Kurniawan, I. T. *et al.* Choosing to make an effort: the role of striatum in signaling physical effort of a chosen action. *J. Neurophysiol.* **104**, 313–21 (2010).
37. Alamia, A., VanRullen, R., Pasqualotto, E., Mouraux, A. & Zénon, A. Pupil-linked arousal responds to unconscious surprisal. *J. Neurosci.* **39**, 3010–18 (2019).
38. Rigoux, L., Stephan, K. E., Friston, K. J. & Daunizeau, J. Bayesian model selection for group studies - Revisited. *Neuroimage* (2014). doi:10.1016/j.neuroimage.2013.08.065
39. Hartmann, M. N., Hager, O. M., Tobler, P. N. & Kaiser, S. Parabolic discounting of monetary rewards by physical effort. *Behav. Processes* **100**, 192–196 (2013).
40. Lockwood, P. L. *et al.* Prosocial apathy for helping others when effort is required. *Nat. Hum. Behav.* **1**, 1–10 (2017).
41. Kiani, R., Corthell, L. & Shadlen, M. N. Choice certainty is informed by both evidence and decision time. *Neuron* (2014). doi:10.1016/j.neuron.2014.12.015
42. Ebitz, R. B. & Platt, M. L. Neuronal activity in primate dorsal anterior cingulate cortex signals task conflict and predicts adjustments in pupil-linked arousal. *Neuron* **85**, 628–640 (2015).
43. Glimcher, P. W. Choice: Towards a Standard Back-pocket Model. in *Neuroeconomics: Decision making and the brain* (eds. Glimcher, P. W., Camerer, C. F., Fehr, E. & Poldrack, R. A.) 503–521 (Academic Press, 2009). doi:10.1016/B978-0-12-374176-9.00032-4
44. Gaesser, B., Keeler, K. & Young, L. Moral imagination: Facilitating prosocial decision-making through scene imagery and theory of mind. *Cognition* **171**, 180–193 (2018).
45. de Gee, J. W., Knapen, T. & Donner, T. H. Decision-related pupil dilation reflects upcoming choice and individual bias. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E618-25 (2014).
46. Schwarz, L. A. *et al.* Viral-genetic tracing of the input-output organization of a central noradrenaline circuit. *Nature* **524**, 88–92 (2015).
47. de Gee, J. W. *et al.* Dynamic modulation of decision biases by brainstem arousal systems. *Elife* **6**, 1–36 (2017).
48. Arnsten, A. F. T. & Goldman-Rakic, P. S. Selective prefrontal cortical projections to the region of the locus coeruleus and raphe nuclei in the rhesus monkey. *Brain Res.* **306**, 9–18 (1984).
49. Dalsass, M., Kiser, S., Mèndershausen, M. & German, D. C. Medial prefrontal cortical projections to the region of the dorsal periventricular catecholamine system. *Neuroscience* (1981). doi:10.1016/0306-4522(81)90149-4
50. Parvizi, J., Rangarajan, V., Shirer, W. R., Desai, N. & Greicius, M. D. Case Study The Will to Persevere Induced by Electrical Stimulation of the Human Cingulate Gyrus. *Neuron* **80**, 1359–1367 (2013).
51. Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nat. Neurosci.* **19**, 1286–1291 (2016).
52. Aston-Jones, G. & Cohen, J. D. An integrative theory of Locus Coeruleus-Norepinephrine function:

712 adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
713 53. Steinmetz, J., Tausen, B. M. & Risen, J. L. Mental Simulation of Visceral States Affects Preferences and
714 Behavior. *Personal. Soc. Psychol. Bull.* **44**, 406–417 (2018).
715 54. Fleming, S. M. & Daw, N. D. Self-evaluation of decision-making: A general Bayesian framework for
716 metacognitive computation. *Psychol. Rev.* **124**, 91–114 (2017).
717 55. Vinckier, F. *et al.* Sour grapes and sweet victories : how actions shape preferences. *PLOS Comput. Biol.*
718 **15**, e1006499 (2018).
719 56. Hassabis, D. & Maguire, E. A. Deconstructing episodic memory with construction. *Trends Cogn. Sci.*
720 (2007). doi:10.1016/j.tics.2007.05.001
721 57. Salamone, J. D., Correa, M., Farrar, A. M. & Mingote, S. Effort-related functions of nucleus accumbens
722 dopamine and associated forebrain circuits. *Psychopharmacology (Berl).* **191**, 461–482 (2007).
723 58. Sara, S. J. & Bouret, S. Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through
724 Arousal. *Neuron* (2012). doi:10.1016/j.neuron.2012.09.011
725 59. Saper, C. B., Fuller, P. M., Pedersen, N. P., Lu, J. & Scammell, T. E. Sleep State Switching. *Neuron* **68**,
726 1023–1042 (2010).
727 60. Kurniawan, I. T., Guitart-Masip, M. & Dolan, R. J. Dopamine and effort-based decision making. *Front.*
728 *Decis. Neurosci.* **5**, 1–10 (2011).
729 61. Meyniel, F. *et al.* A specific role for serotonin in overcoming effort cost. *Elife* **5**, 1–18 (2016).
730 62. Pessiglione, M., Vinckier, F., Bouret, S., Daunizeau, J. & Le Bouc, R. Why not try harder?
731 Computational approach to motivation deficits in neuro-psychiatric diseases. *Brain* **141**, 629–650 (2018).
732 63. Nutt, D. *et al.* The other face of depression, reduced positive affect: The role of catecholamines in
733 causation and cure. *J. Psychopharmacol.* **21**, 461–471 (2007).
734 64. Moret, C. & Briley, M. The importance of norepinephrine in depression. *Neuropsychiatr. Dis. Treat.* **7**, 9–
735 13 (2011).
736
737
738

Acknowledgments

General: The authors thank Yoojin Lee and Zoltan Nagy for assistance in MRI optimisation, Karl Treiber and Miguel A. Garcia for assistance in data collection, and our participants for their voluntary participation. **Funding:** This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 702799 to I.K. and by a grant from the Swiss National Science Foundation SNSF (100019L_173248) to C.C.R..

Author contributions: I.K. and C.C.R. conceived and planned the experiment. I.K. carried out the experiment. I.K. conducted all analyses with input from M.G. and C.C.R.. M.G. provided analytical software. I.K., M.G., and C.C.R discussed and interpreted the results, and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data availability:** All raw and processed data, as well as the code to reproduce all analyses and figures will be made available on github or the OSF upon publication.

Figures and Tables

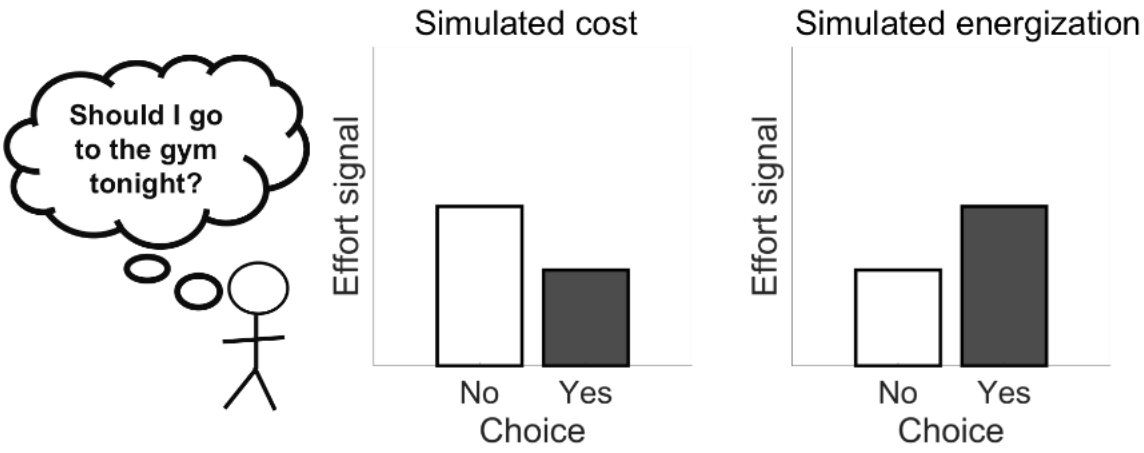

Figure 1. Predicted anticipatory neural response to effort as a function of choice outcomes. "Yes" decisions refer to decisions whereby individuals choose to perform the effort, "No" decisions refer to those whereby individuals decide to forego it.. According to a simulated-cost scenario, effort-related signals should be higher when individuals reject the proposed effort, whereas the simulated-energization scenario predicts that these signals should be higher when individuals accept the proposed effort.

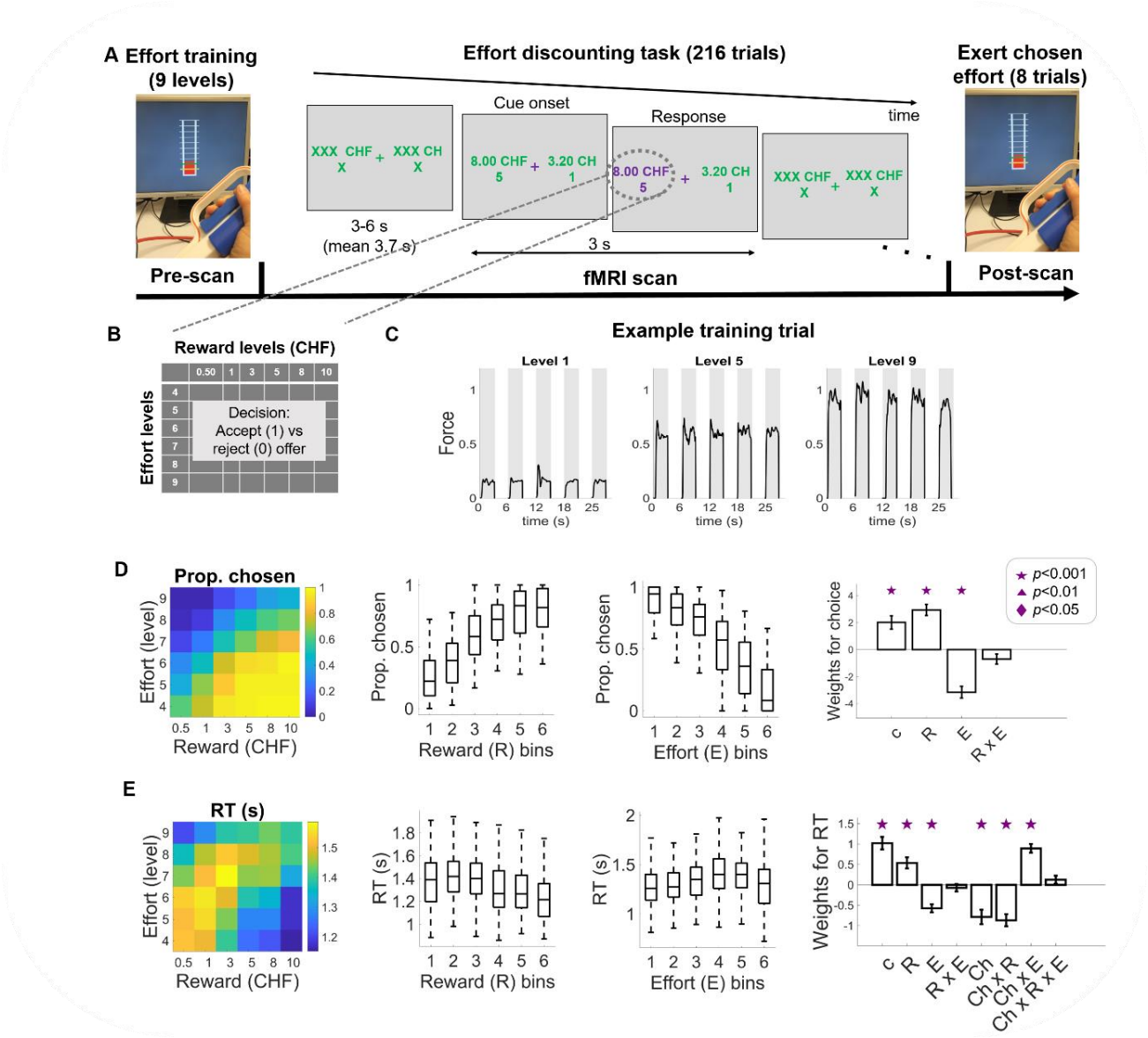

Figure 2. Experimental paradigm and behavioral results. **A)** Pre-scan: Participants received visually-guided effort training on a hand-held dynamometer. Levels 1-9 correspond to 10-90% maximum voluntary contraction (MVC). In fMRI scanner, participants chose between an offer associated with variable amounts of reward and effort and a counteroffer with smaller reward. Post-scan: Outside the scanner, eight randomly selected trials were realized whereby participants executed the effort they chose to obtain the associated reward. **B)** Factorial design of the offer with 6 levels of effort and 6 levels of reward. Reward of the counteroffers (not shown) is either 30% or 40% of the larger offer, and the effort is always the lowest force level (level 1). **C)** Force traces from three example training trials. **D-E)** Behavioral data: Proportions of accepted offers (D) and response times (RT; E) as shown from left to right: color map, main effect of reward, main effect of effort, and multiple regression. Symbols indicate significance levels against zero. Abbreviations: c=regression constant, R=reward levels, E=effort levels, Ch=Choice (1=Accept; 0 reject). Boxplots display the median (central line), 25th and 75th percentiles (bottom and top edges), and non-outlier low and high extreme values (bottom and top error bars). Bar plots display means $\pm$ 1 standard error of the mean (SEM).

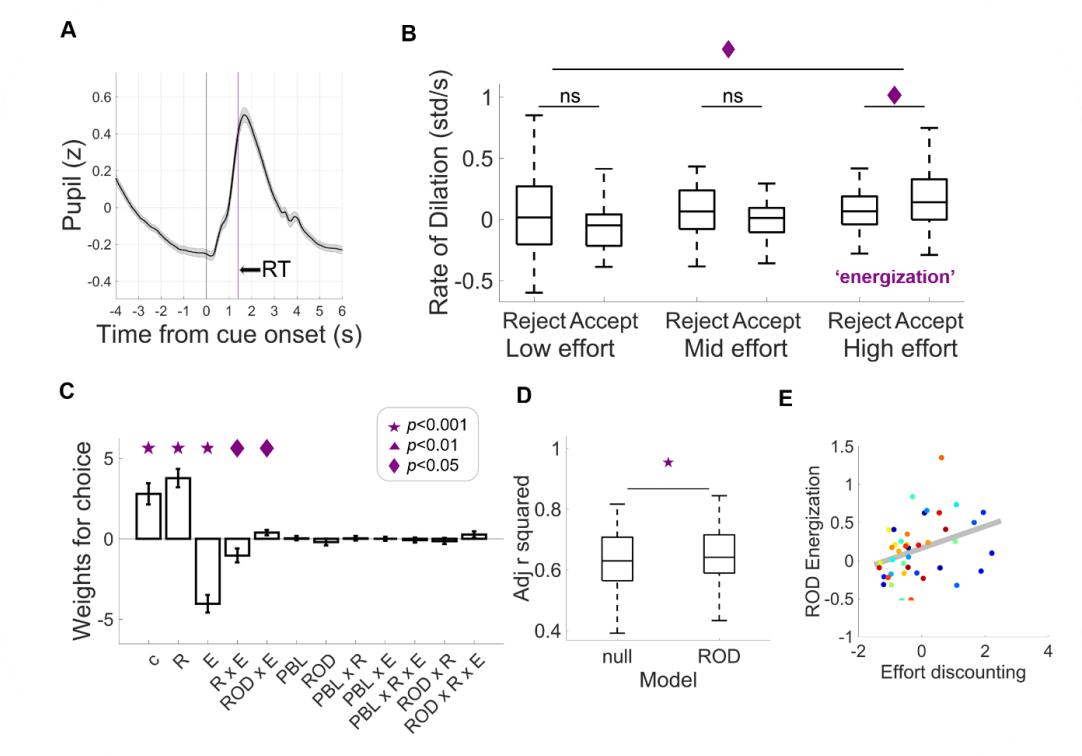

Figure 3. Pupil results. A) Grand-mean of pupillary response time-locked to cue onset. Second vertical line (purple) indicates averaged RT onset. **B)** Significant choice-by-effort interaction effect on rate of pupil dilation (ROD) with choice factor in reject and accept, and effort factor in low, middle, and high bins. **C)** Weights of logistic regression of choice on reward, effort, ROD, PBL, and the interactions. **D)** Adjusted R^2 of the regression model with ROD as shown in figure 3C is significantly higher than that of the null model as shown in figure 2D (right). **E)** Significantly positive correlation between the energization signal in ROD (accept minus reject high effort) and z-scored individual parameter of effort discounting. Each data point represents a subject. All scatterplots use the same color-coding scheme for subjects. Symbols indicate significance levels between indicated conditions (B & D) or against zero (C). Boxplots display the median (central line), 25th and 75th percentiles (bottom and top edges), and non-outlier low and high extreme values (bottom and top error bars). Bar plots display means $\pm$ 1 standard error of the mean (SEM). Abbreviations: c=constant, R=reward levels, E=effort levels, ROD=rate of dilation, PBL=pupil baseline levels.

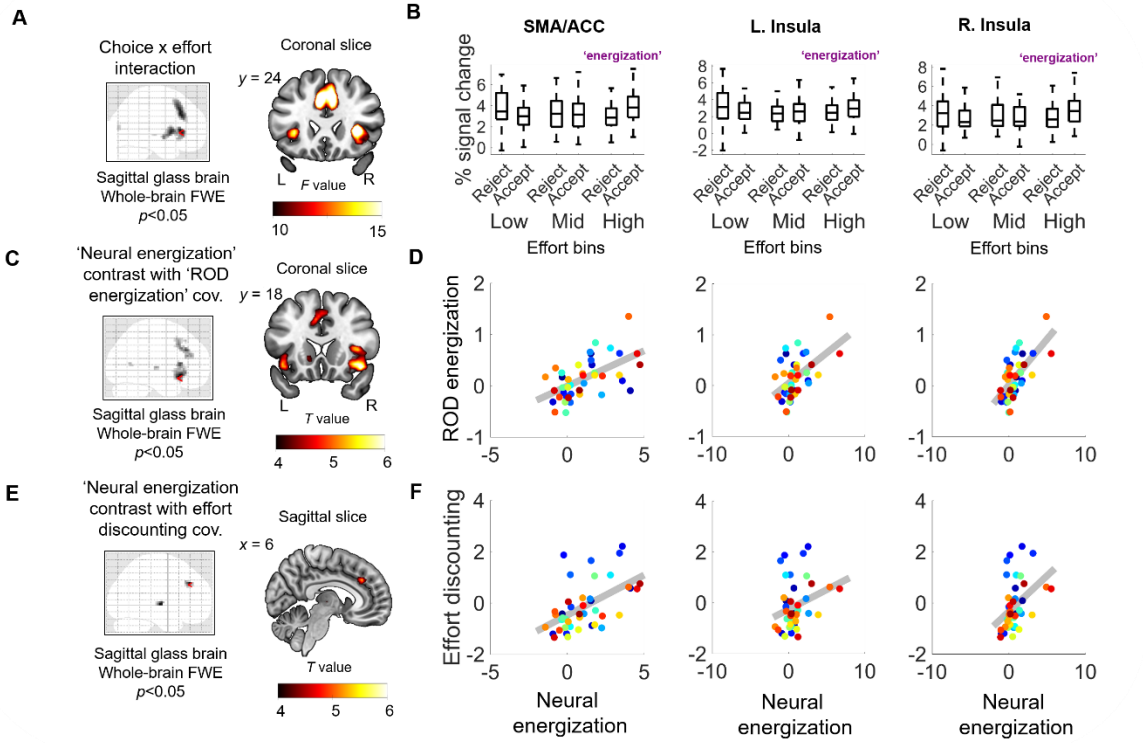

Figure 4. Brain results. SPM of brain activity for cue presentation and correlation with 'ROD energization' and parabolic effort discounting parameter from figure 2C. **A)** Significant choice-by-effort interaction effect in Supplementary Motor Area (SMA)/ dorsal Anterior Cingulate cortex (ACC) and bilateral anterior insula. **B)** For illustration purposes, beta plots of extracted percent BOLD signal change from baseline within all three brain clusters in A as functional ROIs in D & F. **C&E)** Whole brain analysis shows significant correlation between 'neural energization' (accept > reject high effort trials) and subject covariates 'ROD energization' (C) and z-scored effort discounting parameter (E). **D&F)** Similarly, ROI analysis shows significantly positive correlations between 'neural energization' contrasts extracted from all three functional ROIs with 'ROD energization' (D) and effort discounting (F). Each data point represents a subject. All scatterplots use the same color-coding scheme for subjects.

Table 1. MNI coordinates and statistics for choice and effort effects. Here we report main effects of choice and effort, choice-by-effort interaction, and simple effects of choice from GLM1a. Unless otherwise stated, all effects are from t-tests. *P* values are at cluster-level FWE correction.

Effect	Brain region	k	F or t-value	p-value	MNI Coordinates		
					x	y	z
Main effect of choice (F-test)	L Middle Occipital Gyrus	86	28.750	0.002	-42	-72	30
	Location not in atlas	35	25.550	0.025	3	12	-6
	L Inf Parietal Lobule	97	24.030	0.001	-57	-39	30
	L Mid Orbital Gyrus	34	23.690	0.027	-9	42	-9
Accept > reject	L Middle Occipital Gyrus	274	5.362	<0.0001	-42	-72	30
	L SupraMarginal Gyrus	274	4.902	<0.0001	-57	-39	30
	Nucleus accumbens	39	5.055	0.024	3	12	-6
	R Fusiform Gyrus	27	5.009	0.048	42	-33	-12
	L Mid Orbital Gyrus	52	4.867	0.012	-9	42	-9
	R Cerebelum (VI)	51	4.584	0.012	18	-72	-24
	L Middle Frontal Gyrus	56	4.300	0.01	-36	30	39
Main effect of effort (F test)	L Angular Gyrus	97	14.704	0.001	-36	-57	39
	R Angular Gyrus	103	13.837	<0.0001	39	-57	42
High > Mid effort	L Angular Gyrus	249	5.413	<0.0001	-36	-57	39
	R Angular Gyrus	210	4.875	<0.0001	39	-57	42
	R ACC	72	4.800	0.005	9	42	24
	L Precuneus	60	4.316	0.008	-3	-66	33
Choice x Effort (F test)	R Anterior Insula	127	22.094	<0.0001	33	24	3
	L ACC	350	20.748	<0.0001	-9	24	33
	R Caudate Nucleus	123	19.831	<0.0001	12	9	3
	Midbrain	123	13.585	<0.0001	-6	-6	-6
	L Caudate	35	15.200	0.021	-12	6	6
	L Anterior Insula	71	14.765	0.002	-33	24	0
Choice x Effort positive interaction (only high and mid effort)	R MCC	186	5.265	<0.0001	9	24	39

	R Anterior Insula	43	4.598	0.019	30	27	3
	L Calcarine Gyrus	35	4.249	0.03	-6	-84	9
Choice x Effort positive interaction (only high and low effort)	R Insula Lobe	178	6.343	<0.0001	33	24	3
	L ACC	453	6.326	<0.0001	-9	24	33
	L ACC	453	4.230	<0.0001	0	39	21
	R Caudate Nucleus	288	6.250	<0.0001	12	9	3
	L Caudate	288	5.332	<0.0001	-12	6	6
	L IFG (p. Orbitalis)	136	5.433	<0.0001	-33	24	0
Accept > reject high effort	Nucleus accumbens	296	5.615	<0.0001	6	3	-3
	L ACC	302	5.373	<0.0001	-6	24	30
	R ACC	302	4.705	<0.0001	6	36	18
	R Anterior Insula	54	5.241	0.011	27	24	3
	L Superior Frontal Gyrus	39	4.594	0.024	-21	27	54
Accept > reject mid effort	L SupraMarginal Gyrus	176	4.800	<0.0001	-51	-48	33
	L Middle Occipital Gyrus	176	4.448	<0.0001	-42	-72	30
	L Rectal Gyrus	12	4.011	>0.05	-3	42	-12

Table 2. MNI coordinates and statistics for correlations with 'neural energization'. Here we report correlation between 'neural energization' and 'ROD energization' (GLM1b) and between 'neural energization' and effort discounting (GLM1c)**.** Unless otherwise stated, all effects are from t-tests. *P* values are at cluster-level FWE correction.

Effect	Brain region	k	F or t-value	p-value	MNI Coordinates		
					x	y	z
Accept > Reject high effort with accept-reject ROD cov.	R IFG (p. Orbitalis)	203	6.786	<0.0001	42	18	-12
	R ACC	152	5.907	<0.0001	9	30	30
	L MCC	152	5.240	<0.0001	-6	18	39
	L Temporal Pole	61	5.592	0.004	-42	15	-12
	R IFG (p. Triangularis)	26	5.171	0.039	54	33	21
Accept > Reject high effort with effort discounting cov.	Thalamus	26	5.604	0.039	0	-15	0
	R ACC	35	5.410	0.02	6	33	33