Saïd Business School
Research Papers

February 2016

Models of affective decision-making: how do feelings predict choice?

Caroline J. Charpentier
Institute of Cognitive Neuroscience, University College London and Affective Brain Lab, Department of Experimental Psychology, University College London

Jan-Emmanuel De Neve
Saïd Business School, University of Oxford

Jonathan P. Roiser
Institute of Cognitive Neuroscience, University College London

Tali Sharot
Affective Brain Lab, Department of Experimental Psychology, University College London

Saïd Business School RP 2016-08

Models of affective decision-making: how do feelings predict choice?

Running head: Models of affective decision-making

Authors: Caroline J. Charpentier[a,b], Jan-Emmanuel De Neve[c], Jonathan P. Roiser[a], Tali Sharot[b]

Affiliations: [a]Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, UK; [b]Affective Brain Lab, Department of Experimental Psychology, University College London, London WC1H 0AP, UK; [c]Saïd Business School, University of Oxford.

Corresponding authors:

Caroline Charpentier, Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK. Telephone: +44(0)2076791138. E-mail: caroline.charpentier.11@ucl.ac.uk.

Tali Sharot, UCL Psychology, 26 Bedford Way, London WC1H 0AP, UK. Telephone: +44(0)2076797511. E-mail: t.sharot@ucl.ac.uk.

Abstract

Intuitively, how we feel about potential outcomes will determine our decisions. Indeed, one of the most influential theories in psychology, Prospect Theory, implicitly assumes that feelings govern choice. Surprisingly, however, we know very little about the rules by which feelings are transformed into decisions. Here, we characterize a computational model that uses feelings to predict choice. We reveal that this model predicts choice better than existing value-based models, showing a unique contribution of feelings to decisions, over and above value. Similar to Prospect Theory value function, feelings showed diminished sensitivity to outcomes as value increased. However, loss aversion in choice was explained by an asymmetry in how feelings about losses and gains were weighed when making a decision, not by an asymmetry in the feelings themselves. The results provide new insights into how feelings are utilized to reach a decision.

Keywords: decision-making, feelings, subjective well-being, value, utility, Prospect Theory

Introduction

How would you feel if you received international recognition for outstanding professional achievement? How would you feel if your marriage broke apart? Intuitively, answers to these questions are important, as they should predict your actions. If the prospect of losing your spouse does not fill you with negative feelings you may not attempt to keep the unit intact. But how exactly do feelings associated with possible outcomes relate to actual choices? What are the computational rules by which feelings are transformed into decisions? While an expanding body of literature has been dedicated to answering the reverse question, namely how decision outcomes affect feelings (Carter & McBride, 2013; Kassam, Morewedge, Gilbert, & Wilson, 2011; Kermer, Driver-Linn, Wilson, & Gilbert, 2006; McGraw, Larsen, Kahneman, & Schkade, 2010; Mellers, Schwartz, Ho, & Ritov, 1997; Rutledge, Skandali, Dayan, & Dolan, 2014; Yechiam, Telpaz, & Hochman, 2014), little is known of how feelings drive decisions about potential outcomes.

Here, we examine whether feelings predict choice and built a computational model that characterizes this relationship. We turned to Prospect Theory (Fox & Poldrack, 2014; Kahneman & Tversky, 1979; Tversky & Kahneman, 1986, 1992) as a starting point in this endeavor. Prospect Theory was not derived by eliciting people's feelings to predict choice, but rather by observing people's choices in order to estimate the subjective value associated with possible outcomes. An implicit assumption of the theory, however, is that subjective value (utility) is a proxy for feelings, which in turn govern choice; "humans described by Prospect Theory are guided by the immediate emotional impact of gains and losses" (Kahneman, 2011). This suggests that if we measure a person's feelings associated with different outcomes, we should be able to generate that person's utility function and use it to predict their choices. While Prospect Theory is one of the most influential theories in economics and psychology, this implicit assumption has never been empirically tested. Thus, we do not know if and how feelings guide choice.

To address this question, in two separate studies (see Supplemental Material for replication study), participants reported how they felt, or expected to feel, after winning or losing different amounts of money. We used those self-reported feelings to form a "feeling function"; a function that best relates feelings (expected and/or experienced) to objective value. Next, we used this function to predict participants' choices in a different decision-making task. Our findings were replicated in both studies.

An intriguing question is what such a "feeling function" would look like. One possibility is that it resembles Prospect Theory's value function, which relates the subjective value estimated from choice data to objective value. First, for most people, the value function is steeper for losses in comparison to gains. This results in loss aversion, such that the absolute subjective value of losing a dollar is greater than that of winning a dollar. Yet, while losses *appear* to "loom larger than gains" (Kahneman & Tversky, 1979), we do not know whether the impact of a loss on our feelings is greater than the impact of an equivalent gain. Alternatively, it is possible that the impact of gains and losses on feelings is similar, but that the weight given to those feelings differs when making a choice. Second, Prospect Theory's value function is convex in the loss domain while concave in the gain domain (resembling an "S-shape"). The curvature of the function in both domains represents the notion of diminishing sensitivity to changes in value as gains and losses increase. In other words, the subjective value of gaining (or losing) ten dollars is smaller than twice that of gaining (or losing) five dollars. This diminishing sensitivity results in risk aversion in the gain domain and risk seeking in the loss domain, with individuals tending to choose a small sure gain over a high but risky gain, but a high risky loss over a small sure loss. We examined whether our "feeling function" was also concave for gains and convex for losses, implying that similar to value, feelings associated with gains and losses are less sensitive to outcome value as gains and losses increase. That is, the impact of winning (or losing) ten dollars on feelings is less than twice the impact of winning (or losing) five dollars.

Once feelings were modeled using this "feeling function" we asked whether they can predict choice. Understanding how explicit feelings relate to behavior has important real-world implications for domains ranging from policy to industry.

Methods

Subjects. Fifty-nine healthy volunteers were recruited to take part in the experiment via the UCL Subject Pool. Sample size was determined using a power analysis (G*power version 3.1.9.2; Faul, Erdfelder, Lang, & Buchner, 2007). Based on previous studies that have investigated the link between decision outcomes and self-report feelings using within-subjects designs, effect sizes (Cohen's d_z) ranged from .245 to .798, with a mean at .401 (Harinck, Van Dijk, Van Beest, & Mersmann, 2007; Kermer et al., 2006; Yechiam et al., 2014). A sample size of 59 subjects was therefore required to achieve 85% power of detecting an effect size of .401 with an alpha of 0.05. Data collection was therefore stopped after 59 subjects. Three subjects were excluded: one who showed no variation at all in their feelings ratings, one whose data from the gambling task were lost, and one who missed more than 50% of the trials in the gambling task. Final analyses were run on 56 subjects (22 males, mean age 23.9y, age range 19-35y). With 56 subjects included, our post-hoc power to detect a .401 effect size was still 83.8%. All participants gave written informed consent and were paid for their participation. The study was approved by the departmental ethics committee at University College London.

Behavioral tasks. Participants completed two tasks, the order of which was counterbalanced.

1. Feelings Task. In the feelings task, subjects completed 4 blocks of 40 to 48 trials each, in which they reported either expected (Fig. 1A) or experienced (Fig. 1B) feelings associated with a range of wins and losses (between £0.2 and £12), or no change in monetary amount (£0). At the beginning of each trial participants were told how much was at stake and whether it was a win trial (e.g., if you choose the "good" picture, you will win £10) or a loss trial (e.g., if you choose the "bad" picture, you will lose £10). Their task was then to make a simple arbitrary choice between two geometrical shapes, associated with a 50% chance of winning versus not winning (on win trials) or of losing versus not losing (on loss trials). On each trial one novel stimulus was randomly associated with a gain or loss (between £0.2 and £12) and the other novel stimulus with no gain and no loss (£0). Each stimulus was presented once so learning was not possible. There was no way for the participants to know which abstract stimulus was associated with a better outcome. They reported their feelings by answering the questions "How do you feel now?" (experienced feelings, after a choice) or "How do you think you will feel if you win/lose/don't win/don't lose?" (expected feelings, before a choice), using a subjective rating scale ranging from "Extremely unhappy" to "Extremely happy". In 2 of the 4 blocks (counterbalanced order) they reported their expected feelings (Fig. 1A), and in the other 2 blocks, they reported their experienced feelings (Fig 1B). Expected and experienced feelings were collected in different blocks to avoid subjects simply remembering and repeating the same rating. The choice between the two geometrical shapes was simply instrumental and implemented in order to have subjects actively involved with the outcomes.

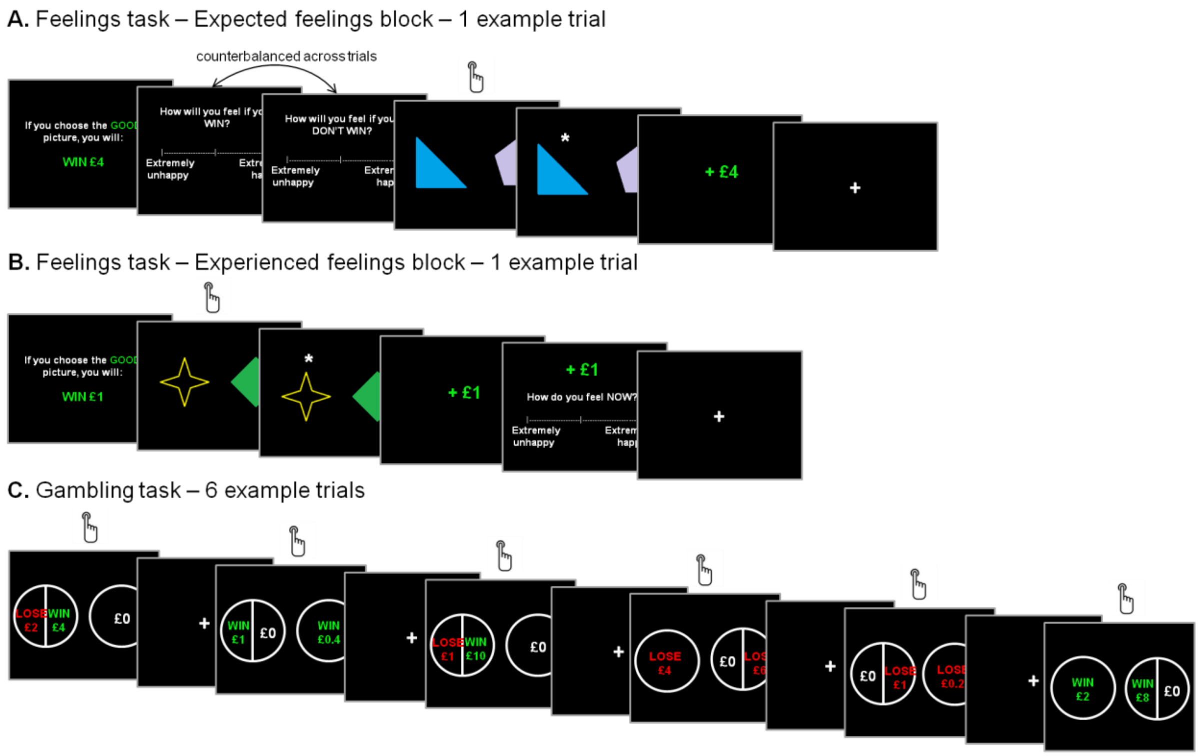

Fig. 1. Experimental design. Participants completed two tasks in a counterbalanced order (**A**,**B**): a feelings task where they reported (in different blocks) expected (**A**) or experienced (**B**) feelings associated with winning, losing, not winning or not losing a range of monetary amounts; and (**C**) a gambling task where they selected between a sure option and a gamble involving the same amounts as those used in the feelings task. Feelings were modeled as a function of value and this resulting feelings function F was used to predict choice in the gambling task. For each trial, feelings associated with the sure option, the risky gain, and the risky loss were extracted and entered in a cross-trials within-subject logistic regression model.

<u>2. Gambling Task.</u> Participants completed a probabilistic choice task (Fig. 1C) in which they made 288-322 choices between a risky 50/50 gamble and a sure option. Importantly, all the amounts used in the gambling task were the same as those used in the feelings task (between £0.2 and £12), such that feelings associated with these outcomes could be combined to predict gamble choice. There were 3 gamble types: mixed (subjects had to choose between a gamble with 50% chance of a gain and 50% of a loss, or sure option of £0), gain-only (subjects had to choose between a gamble with 50% chance of a high gain and 50% chance of £0, or a sure, smaller, gain) and loss-only (subjects had to choose between a gamble with 50% chance of a high loss and 50% chance of £0, or a sure, smaller, loss). In Prospect

Theory, these 3 types of choices are essential to estimate loss aversion, risk preference for gains, and risk preference for losses, respectively.

Subjects started the experiment with an initial endowment of £12 and were paid according to their choices on two randomly chosen trials (across both tasks) at the end of the experiment.

Feelings function models. The impact of outcome on feelings was calculated relative to three different baselines: difference from the mid-point of the rating scale, difference from rating reported on the previous trial (for experienced feelings only), difference from corresponding zero outcome. These were calculated for each win and loss amount, for expected and experienced feelings separately. For each subject, for each of the above methods, feelings function models were then fit (ten for expected feelings and ten for experienced feelings) to explain how feelings best relate to value outcomes:

Feeling Model 1: $$F(x) = \boldsymbol{\beta} x$$

Feeling Model 2: $$F(x) = \begin{cases} \boldsymbol{\beta_{gain}} x, & x > 0 \\ \boldsymbol{\beta_{loss}} x, & x < 0 \end{cases}$$

Feeling Model 3: $$F(x) = \begin{cases} \boldsymbol{\beta}(|x|)^{\boldsymbol{\rho}}, & x > 0 \\ -\boldsymbol{\beta}(|x|)^{\boldsymbol{\rho}}, & x < 0 \end{cases}$$

Feeling Model 4: $$F(x) = \begin{cases} \boldsymbol{\beta_{gain}}(|x|)^{\boldsymbol{\rho}}, & x > 0 \\ -\boldsymbol{\beta_{loss}}(|x|)^{\boldsymbol{\rho}}, & x < 0 \end{cases}$$

Feeling Model 5: $$F(x) = \begin{cases} \boldsymbol{\beta}(|x|)^{\boldsymbol{\rho_{gain}}}, & x > 0 \\ -\boldsymbol{\beta}(|x|)^{\boldsymbol{\rho_{loss}}}, & x < 0 \end{cases}$$

Feeling Model 6: $$F(x) = \begin{cases} \boldsymbol{\beta_{gain}}(|x|)^{\boldsymbol{\rho_{gain}}}, & x > 0 \\ -\boldsymbol{\beta_{loss}}(|x|)^{\boldsymbol{\rho_{loss}}}, & x < 0 \end{cases}$$

Feeling Model 7: $$F(x) = \begin{cases} \boldsymbol{\beta} x + \boldsymbol{\varepsilon}, & x > 0 \\ \boldsymbol{\beta} x - \boldsymbol{\varepsilon}, & x < 0 \end{cases}$$

Feeling Model 8: $$F(x) = \begin{cases} \boldsymbol{\beta_{gain}} x + \boldsymbol{\varepsilon}, & x > 0 \\ \boldsymbol{\beta_{loss}} x - \boldsymbol{\varepsilon}, & x < 0 \end{cases}$$

Feeling Model 9: $$F(x) = \begin{cases} \boldsymbol{\beta} x + \boldsymbol{\varepsilon_{gain}}, & x > 0 \\ \boldsymbol{\beta} x - \boldsymbol{\varepsilon_{loss}}, & x < 0 \end{cases}$$

Feeling Model 10: $$F(x) = \begin{cases} \boldsymbol{\beta_{gain}} x + \boldsymbol{\varepsilon_{gain}}, & x > 0 \\ \boldsymbol{\beta_{loss}} x - \boldsymbol{\varepsilon_{loss}}, & x < 0 \end{cases}$$

In all these models, x represents the value (from -12 to -0.2 for losses and from 0.2 to 12 for gains) and F the associated feeling. The slope between feelings and values is represented by the parameter $\boldsymbol{\beta}$ estimated as a single parameter in all odd-numbered models, or separately for losses and gains in all even-numbered models. If loss aversion is reflected in feelings, $\boldsymbol{\beta_{loss}}$ should be significantly greater than $\boldsymbol{\beta_{gain}}$ and even-numbered models should perform better overall. Similar to the curvature parameter of Prospect Theory value function, $\boldsymbol{\rho}$ reflects the curvature of the feeling function, i.e. the fact that feelings become more or less sensitive to changes in value as absolute value increases (Feeling Models 3 to 6). In Feeling Models 5 and 6, the curvature is estimated separately in the gain and loss domains. If the feeling function is S-shaped (function concave for gains and convex for losses) $\boldsymbol{\rho}$ values should be significantly smaller than 1. To ensure that a function with curvature fit the feelings data better than a simple linear function with an intercept, Feeling Models 7 to 10 were defined (as respective comparisons for Feeling Models 3 to 6), where **ε** represents the intercept, or the offset (positive for gains, negative for losses) where feelings start for values close to £0. All these models were estimated in Matlab (www.mathworks.com) using a maximum-likelihood estimation procedure (Myung, 2003). Bayesian Information Criterion (BIC) were calculated for each subject and model, and then summed across subjects (see Supplemental Material for details). Lower sum of BICs for a given model compared to another indicates better model fit.

Prediction of gambling choice. Feelings values from Feeling Model 3 (found to be the most parsimonious model overall) were then used to predict choices in the gambling task. Specifically, for each participant, the feeling associated with each amount was calculated using Feeling Model 3 with that participant's estimated parameters ($\boldsymbol{\beta}$ and $\boldsymbol{\rho}$). Thus, for each trial of the gambling task, a feelings value was obtained for the sure option, the gain and the loss presented on that trial. A feelings value of 0 was used when the amount in the gamble trial was £0. The probability of choosing the gamble on each trial, coded as 1 if the gamble was chosen and 0 if the sure option was chosen, was then entered as the dependent variable of a logistic regression (Choice Model), with feelings associated with the sure option (S, coded negatively in order to obtain a positive weight), the gain (G, multiplied by its probability 0.5), and the loss (L, multiplied by its probability 0.5) entered as the 3 predictor variables:

$$P(gamble) = \frac{1}{1 + e^{-[\boldsymbol{\omega_S} F(S) + \boldsymbol{\omega_G} F(G) + \boldsymbol{\omega_L} F(L)]}}$$

Logistic regressions were run on Matlab using the glmfit function, using either expected feelings (Choice Model 1) or experienced feelings (Choice Model 2). To determine whether those modeled feelings predicted choice better than value-based models, 5 other comparisons models were used to predict choice from values (Choice Models 3 to 7; see Supplemental Material for details).

In order to be compared across conditions and subjects, weight values ω were standardized using the following equation:

$$\omega'_x = \omega_x \frac{s_x}{s_y}$$

where ω'_x is the standardized weight value, ω_x the original weight for predictor variable x obtained from the regression, s_x the standard deviation of variable x, and s_y the standard deviation of the dependent variable y, here the binary choice values. Standardized weight values were extracted from each regression and compared using repeated-measures ANOVA and paired t-tests.

Replication and extension study. A separate study was conducted to replicate the findings and extend them to cases where the impact of a loss and a gain on feelings is evaluated within the same trial. See Supplemental Material for details and results.

Results

Our analysis followed two main steps. First we used participants' reported feelings associated with different monetary outcomes to build a "feeling function". Specifically, we found the best fitting computational model to characterize how feelings associated with different amounts of gains and losses relate to the objective value of these amounts. Second, we tested whether that model of feelings predicted participants' choices on a separate task. Results of the main study are reported below and results of the replication study in the Supplemental Material.

Characterizing a "feeling function"

Feelings associated with losses and gains were elicited using one of two different scales and the impact of losses and gains on feelings were computed using three different methods (see

Supplemental Material for details): as the change from the mid-point of the rating scale, as the change from the previous rating, and as the change from the rating associated with zero outcome (i.e., the rating associated with not winning or not losing the equivalent amount). For all the models described below the latter baseline resulted in the best fit (Table S1). Thus we report results using this baseline.

We aimed to characterize a model that best fit feelings to outcome value. To that end, for each subject ten models (see Methods for equations) were run to fit data of expected feelings to outcome value and ten equivalent models to fit experienced feelings to outcome value. The models differed from each other in two ways. First, in some models the slope of the function, which indicates how much feelings change for each unit gained/lost, was represented by one parameter (β) and in others by two parameters; one for gains (β_{gain}) and one for losses (β_{loss}). If the latter set of models fit better, that would indicate that gains and losses affect feelings to different extents; if the former set does better that would indicate no difference in the magnitude of influence. Second, models differed with respects to the curvature of the function (ρ). Some models allowed for ρ, some allowed for different curvatures in the loss (ρ_{loss}) and gain (ρ_{gain}) domains, while others did not allow for a curvature at all but rather were linear models with either one intercept (ε) or two intercepts (ε_{gain}, ε_{loss}). If models with a curvature (ρ) fit better than linear models with an intercept (ε) that would suggest that feelings do not increase linearly as a function of outcome value, but that their sensitivity varies as outcomes increase, such that the feeling of winning/losing £10 is more or less intense than twice the feeling of winning/losing £5. Models were estimated using a maximum-likelihood estimation procedure (see Methods for details). Bayesian Information Criterion (BIC), which penalises for additional parameters, showed that the best fitting model (i.e. the lowest BIC value) for both expected (Fig. 2A) and experienced (Fig. 2B) feelings was Feeling Model 3 (see Table S2 for BIC and R^2 values), which has one ρ and one β:

$$F(x) = \begin{cases} \boldsymbol{\beta}(|x|)^{\rho}, & x > 0 \\ -\boldsymbol{\beta}(|x|)^{\rho}, & x < 0 \end{cases} \tag{1}$$

where x is the gain/loss amount (positive for gains and negative for losses) and F the corresponding feeling.

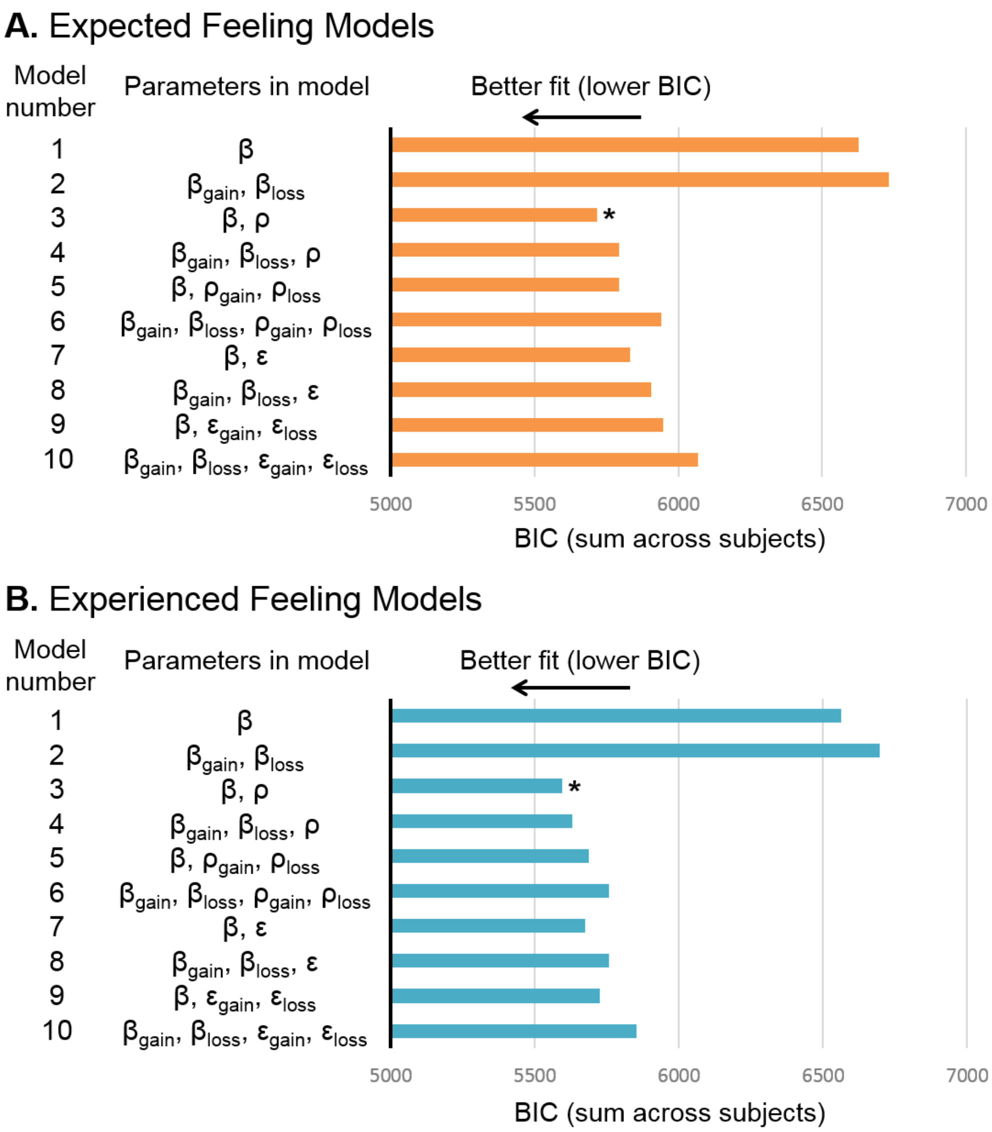

Fig. 2. Feeling Models. BIC values, summed across all subjects, are plotted for ten models fitting feelings to outcome value (see Methods for equations), separately for (**A**) Expected feelings ratings and (**B**) Experienced feelings ratings. Feeling Model 3 was the most parsimonious model, as indicated by lower BIC values for both expected and experienced feelings.

This suggests that:

(i) feelings' sensitivity to outcomes gradually decreased as outcomes increase. Similar to Prospect Theory's value function, ρ was significantly smaller than 1 (expected feelings: ρ=.512 ± SD .26, t(55)=-14.05, *P*<.001, Cohen's d_z=1.88, 95% CI=[.418;.558]; experienced feelings: ρ=.425 ± SD .23, t(55)=-18.52, *P*<.001, Cohen's d_z=2.5, 95% CI=[.513;.637]), indicating that the feeling function was concave in the gain domain and convex in the loss domain. Graphically, we can observe in Fig. 3 that the magnitude of feelings associated with £10 for example was less than twice the magnitude of feelings associated with £5.

(ii) neither sensitivity (β) nor curvature (ρ) differed for gains than losses. Equal sensitivity suggests that when feelings associated with losses and gains are evaluated separately their impact is symmetrical, such that losses are not experienced more intensely than gains. On the surface, these findings contradict the notion of "loss aversion" as proposed by Prospect Theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1986, 1992). However, what we will show later is that while losses do not necessarily impact feelings more than gains they are weighted to a greater extent when making a choice (see Results section on pg 16). With regards to curvature, a single ρ was more parsimonious than two separate ones for gains and losses, suggesting that the extent of concavity for gains was equivalent to the extent of convexity for losses.

Further support for point (i) came from the fact that all models with a curvature parameter ρ (Feeling Models 3-6) were better fits, as indicated by lower BIC values, than corresponding linear models with an intercept (Feeling Models 7-10). This was true both when comparing BICs for models fitting expected feelings (BIC difference < -112) and experienced feelings (BIC difference < -37) (Table S2). Further support for point (ii) came from the fact that Feeling Model 3 had lower BICs than other curved functions with additional parameters that fit gains and losses with separate parameters (Feeling Models 4-6, see Table S3) for both expected and experienced feelings. In addition, the absolute impact of losses and gains on ratings of feelings relative to a zero outcome revealed no difference ($F(1,55)=0.01$, $P=0.92$, $\eta_p^2=.00018$).

Impact bias increases with the amount at stake

Interestingly, comparing the functions for experienced and expected feelings revealed an "impact bias" that increased with amounts lost/gained. The "impact bias" is the tendency to expect losses/gains to impact our feelings more than they actually do (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998). Specifically, the curvature (ρ) was smaller for experienced feeling function relative to expected feeling function (paired t-test: $t(55)=3.31$, $P=0.002$, Cohen's $d_z=.442$, 95% CI=[.034;.138]), while there was no difference in sensitivity values (β) ($t(55)=0.65$, $P=0.52$, Cohen's $d_z=.087$, 95% CI=[-.079;.155]). Thus, although both expected and experienced feelings became less sensitive to outcomes as absolute values of loss/gain increased, this diminished sensitivity was more pronounced in experience than in expectation. As a result, for small amounts of money gained/lost people's expectations of how they will feel were more likely to align with their experience. However, as amounts gained/lost increased, people were more likely to overestimate the effect of outcomes on their feelings, expecting to be affected more by gains and losses than they actually were (i.e., the impact bias (Gilbert et al., 1998)). Graphically, we can observe the growth of the impact bias in Fig. 3 as the increase in separation between the blue line (experienced feelings) and the more extreme orange line (expected feelings).

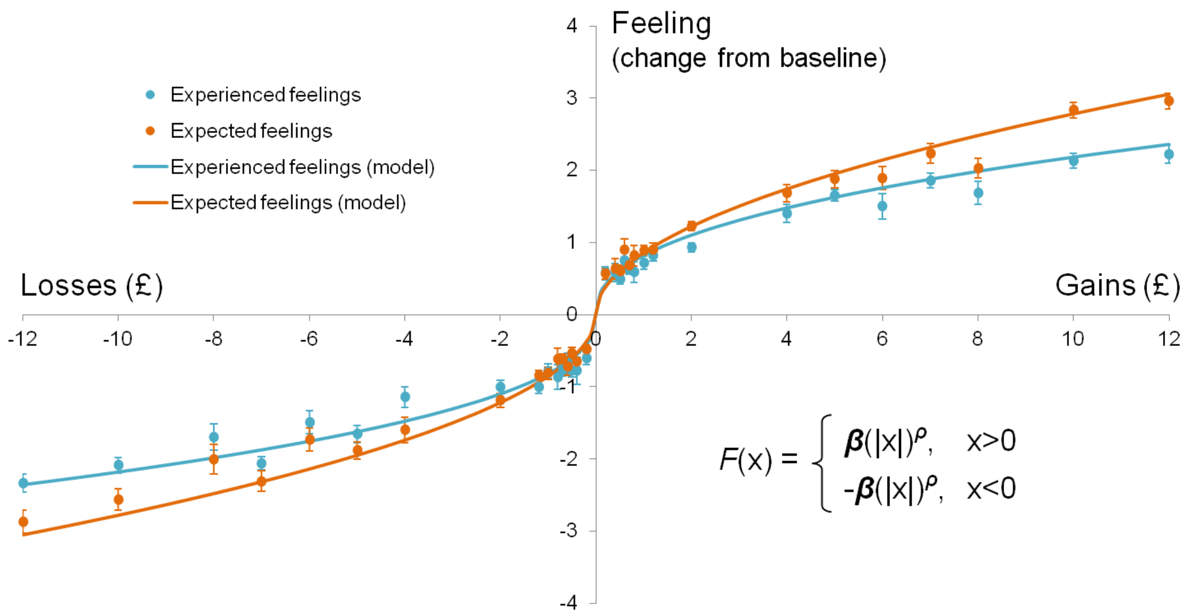

Fig. 3. “Feeling function”. Plotted are expected and experienced feelings ratings averaged across participants for each outcome value, as well as best fitting Feeling Model 3. Average beta (β) across participants, which represents the slope of the function, was 0.857 ± SD 0.36 for expected feelings and 0.819 ± SD 0.37 for experienced feelings (paired t-test revealed no significant difference between them: t(55)=0.65, *P*=0.52, Cohen’s d_z=.087, 95% CI=[-.079;.155]). Average rho (ρ), which represents the curvature of the function, was 0.512 ± SD 0.26 for expected feelings and 0.425 ± SD 0.23 for experienced feelings. Both ρ values were significantly smaller than 1 (t(55)>14, *P*<0.001, Cohen’s d_z>1.87), consistent with an S-shaped function and indicating diminishing sensitivity of feelings to increasing outcome values. ρ was also significantly smaller for experienced relative to expected feelings (paired t-test: t(55)=3.31, *P*=0.002, Cohen’s d_z=.442, 95% CI=[.034;.138]), suggesting that the “impact bias” grows with increasing outcomes. Error bars represent SEM.

Feeling function predicts choice better than value-based models

Once we established a function that fit feelings to outcome value, we turned to the question of how well those feelings predict choices, in particular how they are combined and weighted to make a decision.

To answer this question we used the Feeling Model built above from the data recorded in the first task to predict decisions made in a separate gambling task. To do so we conducted two logistic regressions for each participant (one using expected feelings – Choice Model 1 – and one using experienced feelings – Choice Model 2), where choice on the gambling task was entered as the dependent variable (either 1 if the subject selected the gamble or 0 if the subject selected the sure option) and feelings (predicted by Feeling Model 3) associated with the options were entered as the independent variable. Specifically, using the participant's β and ρ from Feeling Model 3 we computed the feelings associated with each available option multiplied by their probability. For example, if a participant was offered a mixed gamble trial where s/he could either choose a gamble that offered a 50% chance of gaining £10 and a 50% chance of losing £6 or a sure option of £0, we estimated the feelings associated with these three elements multiplied by their probability: the feeling associated with gaining £10 $[F(£10) = \beta \times 10^{\rho} \times 0.5]$; the feeling associated with losing £6 $[F(-£6) = \beta \times (-6)^{\rho} \times 0.5]$ and the feeling associated with getting £0: $[F(£0) = 0 \times 1 = 0]$. These were entered in the logistic regression to predict choice (Choice Model). Each logistic regression thus resulted in three weight parameters ω, which reflected the weight assigned to feelings when making a choice; one for gains (ω_G), one for losses (ω_L) and one for sure options (ω_S).

Importantly, choice models using feelings as predictors (Choice Models 1 and 2) were compared to five other regression models which predicted choice using: objective values (Choice Model 3), log of objective values (consistent with standard economics models to account for the curvature of utility – Choice Model 4), as well as three models derived from Prospect Theory, where value was weighted for each subject with their loss aversion parameter (Choice Model 5), risk aversion parameter (Choice Model 6), or both (Choice Model 7) (see Supplemental Material for more details). To avoid circularity, loss and risk aversion parameters were estimated using half the choice data, and all regression models were tested on the other half.

Feelings, extracted either from the expected or experienced feeling function (Choice Models 1 and 2) predicted choice better than all value-based comparison models (Choice Models 3-

7), as indicated by lower BIC scores (Fig. 4A). Mean R^2 values were also higher for both models predicting choice from feelings (R^2=0.31 for both Choice Models 1 and 2) than for comparison models ($0.26<R^2<0.30$ for Choice Models 3-7), thus consistent with the BIC comparison result. Running the split-half analysis 100 times, with a different way to split the data on every simulation, revealed that models using feelings predicted choice better than all 5 comparison models in 99 simulations out of 100, thus confirming the reliability of this finding.

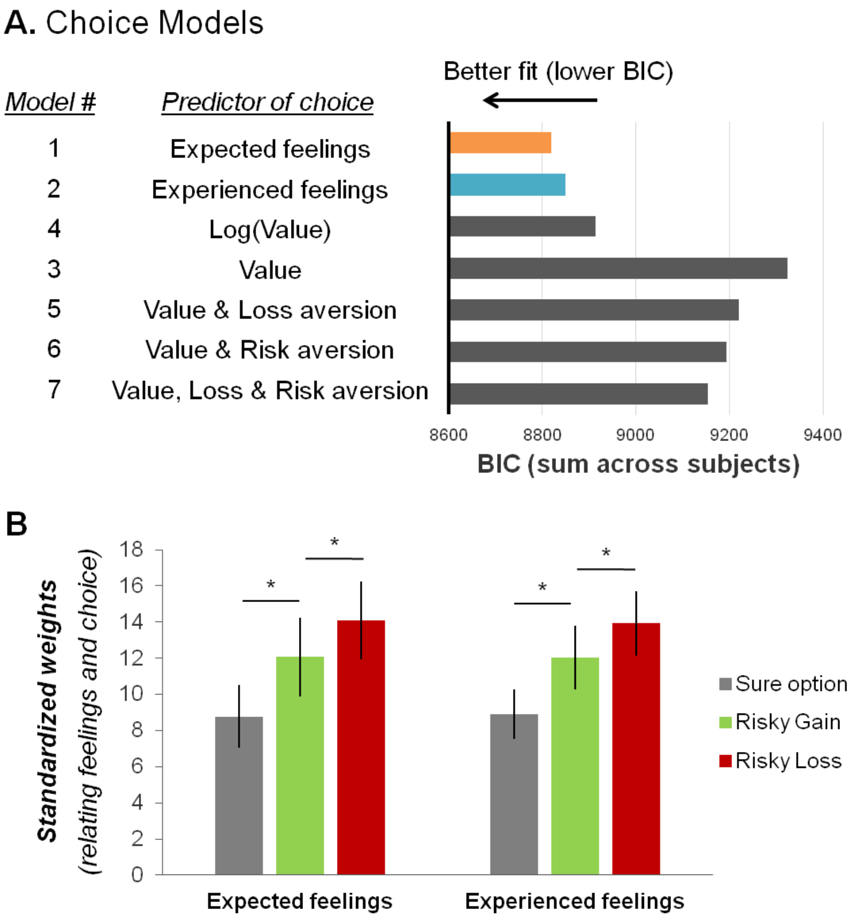

Fig. 4. Choice Models. Seven logistic regressions (or Choice Models) were run to predict choices on the gambling task, using either feelings derived from the "feeling function" build using expected (Choice Model 1) or experienced (Choice Model 2) feelings as predictors, or using value-based comparison models (Choice Models 3-7). (**A**) BIC scores summed across subjects (smaller BIC scores indicate a better fit) show that derived feelings (both expected and experienced) predict choice significantly better than all other value-based models. (**B**) The resulting standardized parameters show that the weight of feelings associated with losses is largest, followed by the weight of feelings associated with gains, with the weight of feelings associated with sure options smallest. This suggests that feelings associated with losses are weighted more than feelings associated with gains. Error bars represent SEM. Two-tailed paired t-tests: * $P<0.05$.

Feelings associated with losses are weighted more than feelings associated with gains when making a decision

Are feelings about potential losses and gains given equal weights when we deliberate on a decision? Our feeling function indicated that the impact of a loss on our feelings was equal to the impact of an equivalent gain. Yet, while losses and gains may impact explicit feelings similarly, we find that these feelings are weighted differently when making a choice.

Specifically, ω parameters from our choice models, which predicted choices from feelings, revealed a greater weight for feelings associated with losses (ω_L) relative to gains (ω_G) in predicting choice (for expected feelings: $t(55)=3.04$, $P=.004$, Cohen's $d_z=.406$, 95% CI=[.684;3.33]; for experienced feelings: $t(55)=2.93$, $P=.005$, Cohen's $d_z=.392$, 95% CI=[.599;3.19]; Fig. 4B). Models that allowed different weights for losses and gains performed significantly better than models that did not (Table S4).

Follow-up analysis revealed that this was true only in mixed-gamble trials, where losses and gains are weighted simultaneously, but not when comparing gain-only and loss-only trials, in which gains and losses are evaluated at different time points (different trials). Specifically, we ran logistic regressions to predict choice from feelings separately for each trial type, and then entered weight of feelings parameters into a two (trial type: mixed/non-mixed) by two (outcome: loss/gain) repeated-measures ANOVA. This revealed a significant interaction (expected feelings: $F(1,55)=6.54$, $P=.013$, $\eta_p^2=.106$; experienced feelings: $F(1,55)=7.46$, $P=.008$, $\eta_p^2=.119$; Fig. S1), driven by a greater weight put on feelings associated with losses relative to gains during mixed-gamble choices (expected feelings: $t(55)=3.66$, $P=.001$, Cohen's $d_z=.489$, 95% CI=[1.67;5.71]; experienced feelings: $t(55)=2.45$, $P=.018$, Cohen's $d_z=.327$, 95% CI=[.91;9.10]) but not during loss- versus gain-only trials (expected feelings: $t(55)=.82$, $P=.42$, Cohen's $d_z=.109$, 95% CI=[-3.25;7.71]; experienced feelings: $t(55)=.79$, $P=.43$, Cohen's $d_z=.105$, 95% CI=[-2.75;6.32]). In other words, only when potential losses and gains are evaluated simultaneously (i.e. in the same gamble) are feelings about losses weighted more strongly during choice than feelings about gains. Results of our replication and extension study supported this claim by showing that even when gains and losses are evaluated in the same trial during the feelings task, their impact on feelings does not differ, but their weight on gamble choice does (see Supplemental Material for details).

To further tease apart the asymmetrical use of feelings associated with gains and losses in shaping choice from the use of value alone, we ran another logistic regression (Choice Model 8, run on all trials regardless of gamble type) in which raw feelings (i.e. reported feelings relative to baseline rather than those derived from the feeling function) were added as predictors of choice in the same logistic regression as objective values themselves. This was done to reveal the weight assigned to feelings in making a choice over and beyond the effect of value *per se*, when the two compete. The results showed no difference in the weight assigned to the value of losses and gains *per se* ($t(55)<1.2$, $P>.23$, Cohen's $d_z<.17$), only to the weight assigned to the associated feelings (expected feelings: $t(55)=3.59$, $P=.001$, Cohen's $d_z=.479$, 95% CI=[1.29;4.55]; experienced feelings: $t(55)=2.28$, $P=.027$, Cohen's $d_z=.307$, 95% CI=[.197;2.89]). Again, this was only true for mixed gamble choices, not for gain-only or loss-only trials where neither feelings nor values were weighted differently between losses and gains (Table S5). This suggests that losses are not weighed differently from gains; rather feelings associated with losses are weighed differently from feelings associated with gains, emphasizing the importance of feelings in decision making.

This last conclusion raises the possibility that individual differences in decision-making could be explained by how people weigh feelings when making a choice. Indeed, using the weights from the above Choice Model 8 we show that individual differences in both loss aversion and the propensity to choose gambles were directly correlated with the extent to which feelings associated with losses were over-weighted compared to gains while controlling for value (correlation between loss aversion and loss-gain weight difference for expected feelings: $r(56)=0.56$, $P<0.001$; for experienced feelings: $r(56)=0.34$, $P=0.012$; correlation between propensity to gamble and loss-gain weight difference for expected feelings: $r(56)=-0.61$, $P<0.001$; for experienced feelings: $r(56)=-0.46$, $P<0.001$; Fig. 5, see Supplementary Information for loss aversion modeling). Specifically, subjects who weighed feelings associated with losses more than gains were more loss averse and less likely to gamble.

This set of results suggests that the asymmetric influence of gains and losses on decision-making, as suggested by Prospect Theory, is neither reflected in expected nor experienced feelings, nor in different weights assigned to value *per se*, but rather in the extent to which feelings associated with losses and gains are taken into account when making a decision.

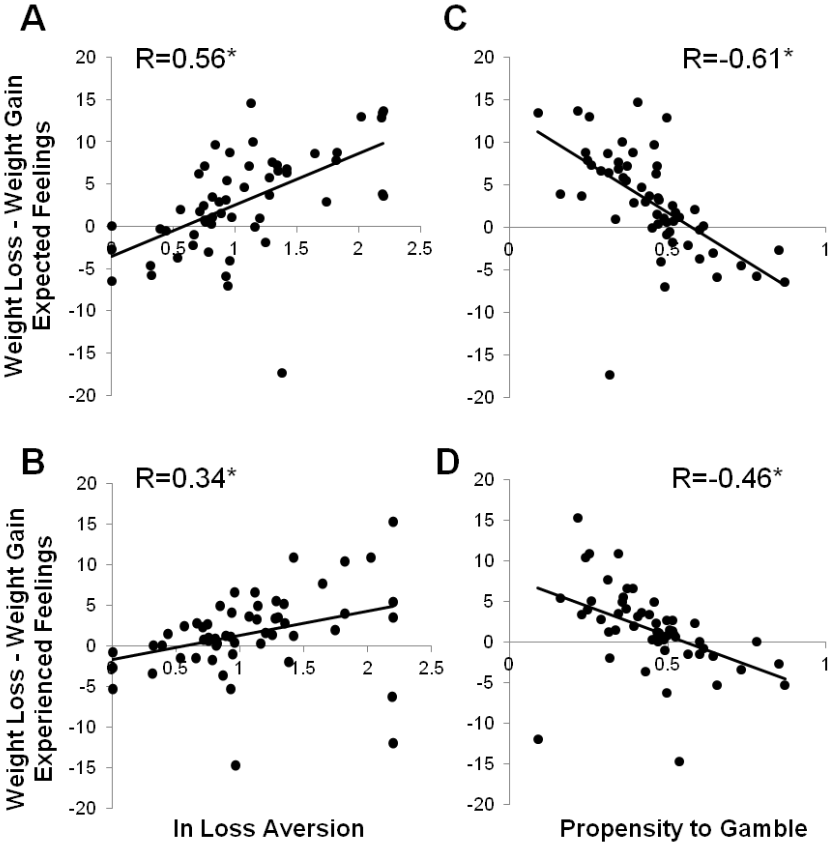

Fig. 5. Individual differences in choice are driven by the relative weights of feelings. Raw feelings (i.e. reported feelings relative to baseline) and objective values were combined in the same regression model (Choice Model 8) to examine the extent to which feelings predict choice while controlling for value. Each regression used either Expected (**A**,**C**) or Experienced (**B**,**D**) raw feelings together with objective values of each of the 3 decision options (Gain, Loss, Sure option), leading to 6 weight parameters in each regression ($\omega_G^{feelings}$, $\omega_L^{feelings}$, $\omega_S^{feelings}$, ω_G^{value}, ω_L^{value}, ω_S^{value}). The difference between the weight of feelings about losses ($\omega_L^{feelings}$) and the weight of feelings about gains ($\omega_G^{feelings}$) was then calculated for each individual and each regression and plotted against ln Loss Aversion (**A**,**B**) (parameter estimated for each individual from the choice data) and proportion of chosen gambles (**C**,**D**). These correlations indicate that the greater weight a participant puts on feelings associated with a loss relative to a gain when making a decision, the more loss averse (and less likely to gamble) they are. Note that loss aversion and propensity to gamble are highly correlated, therefore correlations in **C** and **D** are not independent from **A** and **B**, respectively, and are displayed for illustrations purposes.

Discussion

The relationship between human feelings and the choices they make has occupied scientists, policymakers and philosophers for decades. Indeed, in recent years numerous studies have investigated how decisions and outcomes impact people's feelings (Carter & McBride, 2013; Kassam et al., 2011; Kermer et al., 2006; McGraw et al., 2010; Mellers et al., 1997; Rutledge et al., 2014; Yechiam et al., 2014) and life satisfaction (Boyce, Wood, Banks, Clark, & Brown, 2013; De Neve et al., 2015). Yet, the equally critical question of how people's explicit feelings impact their decisions has been relatively neglected. In this study, we addressed this important question in a controlled laboratory setting and modeled how feelings are integrated into decisions. We demonstrated that feelings drive the decisions people make. However, the rules by which they do so differ from previously assumed.

Feelings were first modeled in a "feeling function" (Feeling Model), which was then used to predict choices (Choice Model). Our Feeling Model predicted choice better than objective values, and a unique contribution of feelings in the decision process was demonstrated. The "feeling function" that best related feelings to value was revealed to be concave for gains and convex for losses, similar to Prospect Theory value function (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) and other non-linear utility functions (Bernoulli, 1954; Fox & Poldrack, 2014; Stauffer, Lak, & Schultz, 2014; Von Neumann & Morgenstern, 1947). This curvature suggests that explicit feelings, similar to subjective value or utility, show diminishing sensitivity to outcomes as the value of these outcomes increases (Carter & McBride, 2013). In other words, the impact of winning or losing ten dollars on feelings is less than twice that of winning or losing five dollars.

Our Feeling Model also revealed no asymmetry between gains and losses, suggesting that the impact of a loss on feelings is not necessarily greater than the impact of an equivalent gain. This was replicated in a separate study extending the symmetrical impact of gains and losses on feelings to cases where a gain and a loss have to be evaluated at the same time. Nevertheless, loss aversion was still present in choice, consistent with Prospect Theory. Importantly, when making a decision a greater weight was put on feelings associated with losses relative to gains. This finding suggests that losses may not impact feelings more strongly than gains as previously implied, but rather that feelings about losses are weighted more when making a choice than feelings about gains. Moreover, the amount by which feelings associated with losses are over-weighted relative to gains in making a decision relates to individual differences in loss aversion and propensity to gamble.

This finding resolves a long-standing puzzle by which loss aversion is often observed in choice, but not necessary in explicit feelings (Harinck et al., 2007; Kermer et al., 2006; McGraw et al., 2010; Mellers et al., 1997). We suggest that the asymmetric influence of gains and losses on decision making, as suggested by Prospect Theory, is not reflected in expected or experienced feelings directly, neither in different weights assigned to value *per se*, but in the extent to which feelings about losses and gains are taken into account when making a decision. Our result is consistent with the interpretation of an increased attention to losses (Yechiam & Hochman, 2013). When losses and gains are presented separately they are experienced in a symmetrical way. However, when they compete for attention, as is the case in the mixed gambles, people may allocate more attention to the feelings they would derive from the loss than from the gain, leading them to choose in a loss averse manner. Another possibility is that people implicitly experience losses to a greater extent than gains (Hochman & Yechiam, 2011; Sokol-Hessner et al., 2009), but this difference is not exhibited in explicit reports.

Our findings also provide the first demonstration of an increasing impact bias with value. Specifically, we found evidence for a general impact bias in feelings (also called affective forecasting error), where people expect the emotional impact of an event to be greater than their actual experience (Gilbert et al., 1998; Kermer et al., 2006; Kwong, Wong, & Tang, 2013; Levine, Lench, Kaplan, & Safer, 2013; Morewedge & Buechel, 2013; Wilson & Gilbert, 2013). Interestingly, this impact bias was not constant, but increased with value. This was due to a stronger curvature of experienced feelings relative to expected feelings. In other words, as absolute value increases, sensitivity to value diminished more quickly for experienced relative to expected feelings. This suggests that as people win or lose more money, they are more and more biased towards overestimating the emotional impact of these outcomes.

Our modeling approach provides novel insight into how explicit feelings relate to choice. Such understanding is both of theoretical importance and has practical implications for policy-makers, economists and clinicians who often measure explicit feelings to predict choice (Benjamin, Heffetz, Kimball, & Rees-Jones, 2012, 2014).

Authors' contributions

C.J. Charpentier, J-E. De Neve, and T. Sharot developed the study concept and design. C.J. Charpentier performed data collection and analysis. C.J. Charpentier and T. Sharot drafted the manuscript. All authors discussed data analysis and interpretation, provided critical revisions, and approved the final version of the manuscript for submission.

Acknowledgments

This work was supported by a UCL Grand Challenge award to C.J.C., a Wellcome Trust Career Development Fellowship to T.S, and a US National Institute on Aging (Grant R01AG040640) and UK Economic & Social Research Council grant to JE.D. We thank Benedetto De Martino, Ori Heffetz, Neil Garrett, Sebastian Bobadilla-Suarez, Stephanie Lazzaro, and Vincent Valton for helpful comments on a previous version of this manuscript; Robb Rutledge, Dan Benjamin, Miles Kimball, and Richard Layard for helpful discussion and suggestions; and Eleanor Barley and Joshua Lim for assistance with data collection.

References

Benjamin, D. J., Heffetz, O., Kimball, M. S., & Rees-jones, A. (2012). What do you think would make you happier? What do you think you would choose? *Am Econ Rev*, *102*(5), 2083–2110.

Benjamin, D. J., Heffetz, O., Kimball, M. S., & Rees-jones, A. (2014). Can marginal rates of substitution be inferred from happiness data? Evidence from residency shoices. *Am Econ Rev*, *104*(11), 3498–3528.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, *22*(1), 23–36.

Boyce, C. J., Wood, A. M., Banks, J., Clark, A. E., & Brown, G. D. A. (2013). Money, well-being, and loss aversion: does an income loss have a greater effect on well-being than an equivalent income gain? *Psychol Sci*, *24*(12), 2557–2562. http://doi.org/10.1177/0956797613496436

Carter, S., & McBride, M. (2013). Experienced utility versus decision utility: Putting the "S" in satisfaction. *J Socio-Econ*, *42*, 13–23. http://doi.org/10.1016/j.socec.2012.11.009

De Neve, J.-E., De Keulenaer, F., Kavetsos, G., Norton, M., Van Landeghem, B., & Ward, G. (2015). Individual experience of positive and negative growth is asymmetric: global evidence using subjective well-being data. *Centre for Economic Performance Discussion Paper No 1304*.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fox, C. R., & Poldrack, R. A. (2014). Prospect Theory and the brain. In *Handbook of Neuroeconomics (2nd edition)*. New York: Elsevier.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: a source of durability bias in affective forecasting. *J Pers Soc Psychol*, *75*(3), 617–638.

Harinck, F., Van Dijk, E., Van Beest, I., & Mersmann, P. (2007). When gains loom larger than losses: reversed loss aversion for small amounts of money. *Psychol Sci*, *18*(12), 1099–1105. http://doi.org/10.1111/j.1467-9280.2007.02031.x

Hochman, G., & Yechiam, E. (2011). Loss aversion in the eye and in the heart: the autonomic nervous system's responses to losses. *J Behav Decis Mak*, *24*(2), 140–156. http://doi.org/10.1002/bdm

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: an analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Kassam, K. S., Morewedge, C. K., Gilbert, D. T., & Wilson, T. D. (2011). Winners love winning and losers love money. *Psychol Sci*, *22*(5), 602–606. http://doi.org/10.1177/0956797611405681

Kermer, D. A., Driver-Linn, E., Wilson, T. D., & Gilbert, D. T. (2006). Loss aversion is an affective forecasting error. *Psychol Sci*, *17*(8), 649–653.

Kwong, J. Y. Y., Wong, K. F. E., & Tang, S. K. Y. (2013). Comparing predicted and actual affective responses to process versus outcome: and emotion-as-feedback perspective. *Cognition*, *129*(1), 42–50.

Levine, L. J., Lench, H. C., Kaplan, R. L., & Safer, M. A. (2013). Like Schrodinger's cat, the impact bias is both dead and alive: reply to Wilson and Gilbert (2013). *J Pers Soc Psychol*, *105*(5), 749–756. http://doi.org/10.1037/a0034340

McGraw, A. P., Larsen, J. T., Kahneman, D., & Schkade, D. (2010). Comparing gains and losses. *Psychol Sci*, *8*(6), 423–429. http://doi.org/10.1177/0956797610381504

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: emotional reactions to the outcomes of risky options. *Psychol Sci*, *8*(6), 423–429.

Morewedge, C. K., & Buechel, E. C. (2013). Motivated underpinning of the impact bias in affective forecasts. *Emotion*, *13*(6), 1023–1029. http://doi.org/10.1037/a0033797

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*(1), 90–100. http://doi.org/10.1016/S0022-2496(02)00028-7

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proc Natl Acad Sci U S A*, *111*(33), 12252–12257. http://doi.org/10.1073/pnas.1407535111

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proc Natl Acad Sci U S A*, *106*(13), 5035–5040.

Stauffer, W. R., Lak, A., & Schultz, W. (2014). Dopamine reward prediction error responses reflect marginal utility. *Curr Biol*, *24*(21), 2491–2500. http://doi.org/10.1016/j.cub.2014.08.064

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *J Bus*, *59*(4), S251–S278.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: cumulative representation of uncertainty. *J Risk Uncertainty*, *5*(4), 297–323.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior* (2nd ed.). Princeton, NJ: Princeton University Press.

Wilson, T. D., & Gilbert, D. T. (2013). The impact bias is alive and well. *J Pers Soc Psychol*, *105*(5), 740–748. http://doi.org/10.1037/a0032662

Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: review and analysis of the unique effects of losses over gains. *Psychol Bull*, *139*(2), 497–518. http://doi.org/10.1037/a0029383

Yechiam, E., Telpaz, A., & Hochman, G. (2014). The complaint bias in subjective evaluations of incentives. *Decision*, *1*(2), 147–160. http://doi.org/10.1037/dec0000008