bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Does endogenous variation in stress modulate risk and time preferences?

Evgeniya Lukinova[1,2,*], Jeffrey C. Erlich[1,2,3,*]

1 NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, China
2 NYU Shanghai, Shanghai, China
3 Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), East China Normal University, Shanghai, China

*** jlukinova@nyu.edu, jerlich@nyu.edu**

Abstract

It has been argued that one dimension of the cycle of poverty is that poverty is a state of chronic stress and that chronic stress impairs decision-making. These poor decisions, made under chronic stress, might include carrying high-interest loans, failure to buy health insurance, gambling or drug use. As such, these decisions can contribute to the cycle of poverty. More specifically, a few studies suggest that increased stress may lead to more risk-aversion and steeper delay-discounting. While the deleterious effects of chronic stress on brain function are well established, much less is known about how chronic stress influences financial decision making specifically. Here, in a longitudinal design within six weeks period we aimed to incorporate biological mechanisms to improve our understanding of how stress influences economic decisions. We used a combination of decision-making tasks, questionnaires, saliva and hair samples within-subject (N=41). We assessed time and risk preferences using hierarchical Bayesian techniques to both pool data and allow heterogeneity in decision making and compared those to cortisol levels and self-reported stress. We found only weak links between endogenous variation in stress and model-based estimates of risk and time preferences. In particular, we found that fluctuations in the stress level measured via hair sample were not only positively correlated with time preferences in the short delay task and risk preferences, but also the decision noise in the risk task. However, relationships for the risk task disappeared when an outlier was removed. Also, we found model-free task measures in the short delay task to be moderately related to both hair cortisol as well as the stressful life events questionnaire measure. For example, we observed that endogenous stress fluctuations and the life change units were negatively correlated with the proportion of later choices. Finally, we established that for the reaction times the curvilinear relationship was preferred to the linear one for those with increase in biological stress level compared to the baseline: when cortisol level increased slightly, participants decided slower, but when stress increased to higher levels, they decided quicker.

Introduction

Stress is known to have negative influence on physical health (Harvey et al., 2003; Association and others, 2010) and psychological well-being (Faravelli and Pallanti, 1989; Hammen, 2005; Kendler et al., 1999; Kessing et al., 2003). Unfavorable effects of stress on behavior can be seen through overeating or undereating, angry outbursts, and social withdrawal. There are a few dominant theories of how stress impacts the quality of decision making (Kandasamy et al., 2014; Haushofer and Fehr, 2014; Mullainathan and Shafir, 2013; Cohn et al., 2015; Lerner et al., 2012; Cornelisse et al., 2013). Still, many studies find no significant effect of acute stress on economic choices (Sokol-Hessner et al., 2016; Lempert et al., 2012). Kandasamy et al. (2014) show that cortisol tablets administered for over 1 week significantly increased risk aversion, whereas Cornelisse et al. (2013) find that continuous administration of hydrocortisone significantly increased the number of impatient choices. Studies in the field also show that chronic stress correlates with poor economic decisions. For example,

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

chronically-stressed individuals carry high-interest loans, have higher incidences of substance abuse, display a lower willingness to take risks and forgo current income for larger future incomes (Haushofer and Fehr, 2014; Mullainathan and Shafir, 2013). However, all of these phenomena are tied to socioeconomic status, making it very challenging to infer causal mechanisms. Although risk aversion typically is seen as benign, recent research documents that people living in poverty are more risk averse and more likely to discount future payoffs (Haushofer and Fehr, 2014; Ceccato et al., 2015; Cueva et al., 2015). Inadequate access to resources may also significantly lower cognitive function reducing our ability to solve problems and to regulate our cognitive activities and, therefore, may lead to stress (Shah et al., 2012). The resulting behavior may contribute to the "cycle of poverty": where children born into poverty tend to stay in poverty (Haushofer and Fehr, 2014). More so, stress may cause maladaptive and dangerous behaviors, making this study relevant to the prevention of substance dependence and various impulsivity disorders (Linsky and Straus, 1986; Lempert et al., 2012, 2018b; Lopez-Guzman et al., 2019).

In our recent studies, we found that students' delay discounting for offers in seconds and days was strongly correlated and stable (Lukinova et al., 2019; Lukinova and Erlich, 2021). This longitudinal within-subject study builds on our previous results about stability of time preferences and explores a biological basis for changes in time preferences over time that may be due to individual responses to the environment (e.g., stressful negative income shocks or non-economic stressors, like divorce, health problems, etc.) using a general population sample ($N = 41$). We define stress as the "organism's reaction to environmental demands exceeding its regulatory capacity" (Haushofer and Fehr, 2014). When a stressor is perceived, the hypothalamic-pituitary-adrenal-axis (HPA-axis) is activated and it releases (among other hormones) cortisol (Schepers and Markus, 2015). Called "the stress hormone", cortisol influences, regulates or modulates many of the changes that occur in the body in response to stress. In healthy individuals, cortisol levels peak in the early morning, and gradually decrease thereafter. To help overcome this challenge, most studies obtain multiple human specimen samples from the time of waking up until sleep or collect samples at the same time for several days.

Several common measures of chronic stress used in the recent literature include questionnaires (Holmes and Rahe, 1967; Goodman et al., 1998), collecting heart rate and blood pressure (Gooding et al., 2015), hair (Ceccato et al., 2015; Karlén et al., 2011), and saliva samples (Pruessner et al., 2003; Inder et al., 2012; Kandasamy et al., 2014). Stress questionnaires can track self-reported 'stress' traits or measure perceived stress. Saliva samples provide a measurement of the cortisol concentration at a single point in time, thus, being more useful to test acute changes, rather than long-term systemic cortisol exposure (Hellhammer et al., 2009). Instead, compared to all other biological stress measures hair samples have certain advantages, including its ability to be stored and transported at room temperature and to reflect cortisol levels over the period of hair growth (one month cortisol level per 1 cm of hair) that can possibly correspond to chronic, rather than acute stress (Karlén et al., 2011). Still, instead of tracking the peak of the diurnal cycle of cortisol release using saliva measure (Inder et al., 2012; Dietrich et al., 2013; Meggs et al., 2016), it is also possible to track the level of cortisol over several days at the same exact time (Pruessner et al., 2003; González-Cabrera et al., 2014). Importantly, when saliva cortisol is aggregated across several time points (after the peak) or several days within a month it is moderately correlated with hair cortisol (Weckesser et al., 2019). Therefore, saliva cortisol measured from a sample taken at the same day of the week and time and compared between these measures might also represent the average change in chronic cortisol level (Lenow et al., 2017). All measures of stress are prone to be collected with some error. Having several measurements in this study and being able to extract a few, allows for better control and validation of the measurement system.

Here, we addressed the limitations of the previous literature and used a battery of decision-making tests and stress measures in order to find out whether chronic stress and its variation account for any variability in choice. To our knowledge, this is the first attempt to bring together intertemporal and risk tasks, hair and saliva bio stress markers and self-reported stress measures in a within-subject longitudinal study using a sample of healthy working adults. First, we assessed stability of time preferences across time horizons and their relation to risk preferences. Then, we explored similarity of distinct stress measures via correlation and principal component analysis. Finally, using stress measures together with behavioral results (both model-free and model-based) allowed us to improve the scholarship on how time and risk preferences interact with hormonal and neuromodulatory systems.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Results

Subjects were recruited from the general population of Shanghai, China (n=41). They participated in a battery of decision-making tasks and questionnaires, total three sessions bi-weekly (Figure 1A). All three sessions were scheduled at approximately the same time for each participant. We collected one saliva sample per participant per session and one hair sample per participant two weeks after session 3.

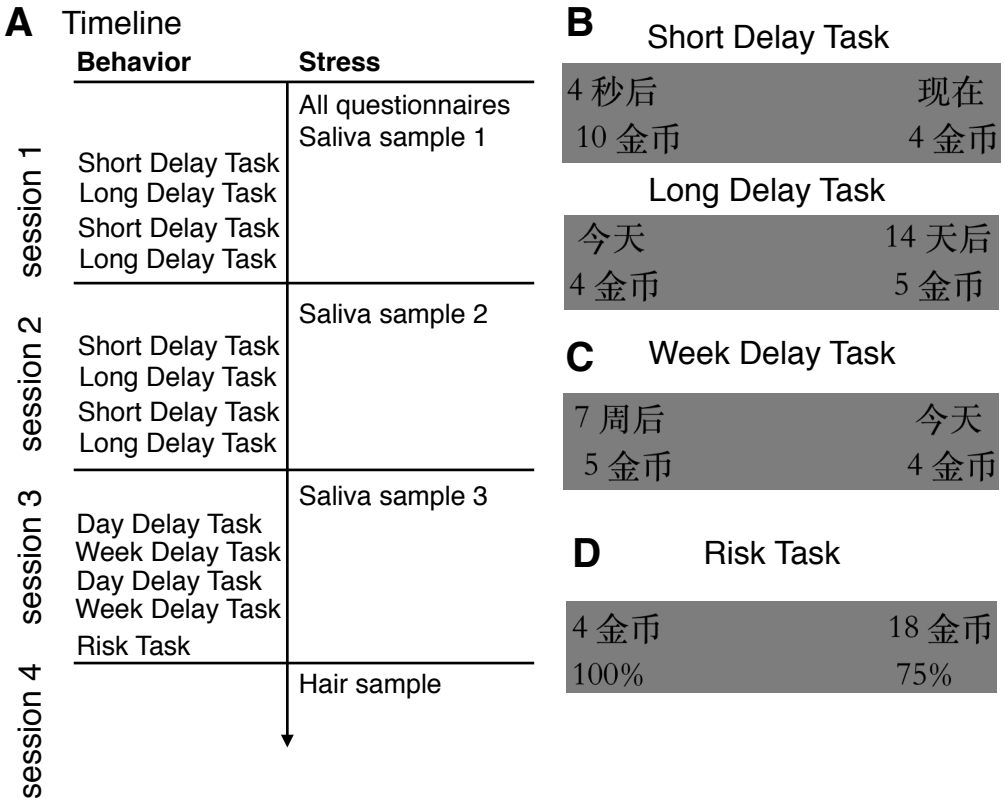

Figure 1. Behavioral Tasks. All stimuli were in Chinese. (A) Timeline of experimental sessions and tasks. (B) Stimuli examples in the short delay (left choice: "in 4 seconds, 10 gold coins", right choice: "now, 4 gold coins") and long delay tasks (left choice: "today, 4 gold coins", right choice: "in 14 days, 5 gold coins"). (C) Stimuli example in the week delay task (left choice: "in 7 weeks, 5 gold coins", right choice: "today, 4 gold coins"). The stimuli for the day delay task were the same as for the long delay task. (D) Stimuli example in the risk task (left choice: "4 gold coins, 100%", right choice: "18 gold coins, 75%").

In each session, subjects completed a series of intertemporal choice tasks. During the first two sessions, subjects participated in tasks with delays in days and seconds (200 trials per task, Figure 1B). During the third session, subjects participated in tasks with delays in days and weeks (100 trials per task, Figure 1C). In each trial of delay-discounting task, subjects made a decision between the sooner and the later options. Only in the short delay verbal task (delays in seconds), subjects experienced the delay whenever delayed option was chosen at the end of each trial. Also, rewards from each trial were accumulated. In contrast, for other delay-discounting tasks (delays in days or in weeks), subjects saw their choice confirmation and then proceeded to the next trial. At the end of the session, a single trial was selected randomly to determine the payment. To summarize, across three experimental sessions subjects participated in three distinct tasks: short delay verbal task (delays in seconds, SV), long delay verbal task (LV, which was exactly the same as day delay verbal task DV = LV) with delays in days, and week delay verbal task (delays in weeks, WV).

During the third session, after the delay-discounting tasks subjects also participated in the risk task (200 trials divided into two equal parts with a break in between, Figure 1D). On each trial the choice was made

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1. Correlations of subjects' discount factors

	Reduced Model	Noise Per Unit Model	Reward Scale Model	N
between tasks				
sessions 1 & 2 (SV vs. LV)	$r = 0.49$	$r = 0.50$	$r = 0.46$	41
	$p = 0.0010$	$p = 0.0009$	$p = 0.0026$	
session 1 (SV1 vs. LV1)	$r = 0.52$	$r = 0.53$	$r = 0.50$	41
	$p = 0.0005$	$p = 0.0004$	$p = 0.0008$	
session 2 (SV2 vs. LV2)	$r = 0.47$	$r = 0.50$	$r = 0.48$	40
	$p = 0.0022$	$p = 0.0010$	$p = 0.0019$	
session 3 (WV vs. DV (LV3))	$r = 0.97$	$r = 0.97$	$r = 0.96$	41
	$p < 2.2e^{-16}$	$p < 2.2e^{-16}$	$p < 2.2e^{-16}$	
between sessions				
SV1 vs. SV2	$r = 0.60$	$r = 0.60$	$r = 0.66$	40
	$p = 4.684e^{-05}$	$p = 4.882e^{-05}$	$p = 3.96e^{-06}$	
LV1 vs. LV2	$r = 0.53$	$r = 0.56$	$r = 0.52$	40
	$p = 0.0004$	$p = 0.0002$	$p = 0.0006$	
LV2 vs. LV3	$r = 0.89$	$r = 0.88$	$r = 0.87$	40
	$p = 2.374e^{-14}$	$p = 1.526e^{-13}$	$p = 2.974e^{-13}$	
LV1 vs. LV3	$r = 0.54$	$r = 0.54$	$r = 0.54$	41
	$p = 0.0002$	$p = 0.0002$	$p = 0.0003$	

Note. Fits for the main 'reward scale' model were done with a hierarchical Bayesian model and had six population level parameters (log discount factor, $\log(k)$; decision noise, $\log(\tau)$; and reward scale, rews; for each of the two delay-discounting tasks per session or per group of sessions) and three parameters per subject ($\log(k_{SV})$, $\log(k_{LV})$, $\log(\tau)$). In session 2, one subject was not sensitive to delay and was excluded from that session fit (reflected in N).

between a risky option (a probabilistic reward) and a safe option (surebet). The was no feedback (on whether the lottery was won or lost) during the Risk Task: after the choice subjects saw a confirmation of their choice and proceeded to the next trial. After the task finished, two random trials were selected (one from each part), the lottery realised and subject was paid.

Stability of time preferences across time and sessions

Subjects' time preferences were estimated by fitting their choices with a Bayesian hierarchical model (BHM) of hyperbolic discounting with decision noise and reward scaling (Materials and Methods). The delay-discounting fits were done jointly for the first two sessions (DS, delays in days and seconds; 16,182 choices), split into two consecutive DS sessions, and separately for the last session (DW, delays in days and weeks; 8,095 choices) across 41 subjects in three delay-discounting tasks: SV, LV (=DV), and WV. The subjects' choices were well fit by the model as assessed using Bayesian r^2 and reported in each of the subject delay plots (Figure 2 panels E, F for example subjects or see SI for all subjects plots).

We find that time preferences are stable across time-horizons: between days and seconds, $r = 0.46$ and between days and weeks, $r = 0.96$ (Figure 2A-D). Therefore, we are able to replicate our previous results (Lukinova et al., 2019; Lukinova and Erlich, 2021) in a nonstudent population. We also report the between tasks and between sessions Pearson correlations of subjects' discount factors ($\log(k)$) in Table 1 for all models under consideration. The correlation coefficient between SV and LV tasks decreases with the increase in number of parameters in the model (nonsignificant difference according to 'cocor' tests). The between session correlation decreases slightly when reward scaling is used only between first and second sessions of the LV task.

The main model was compared to two additional models that reflected the reduced number of parameters. Based on 10-fold cross validation criteria all comparisons favored the 'reward scale' model (Table 2). We further used the 'reward scale' model (the main model) results in our analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

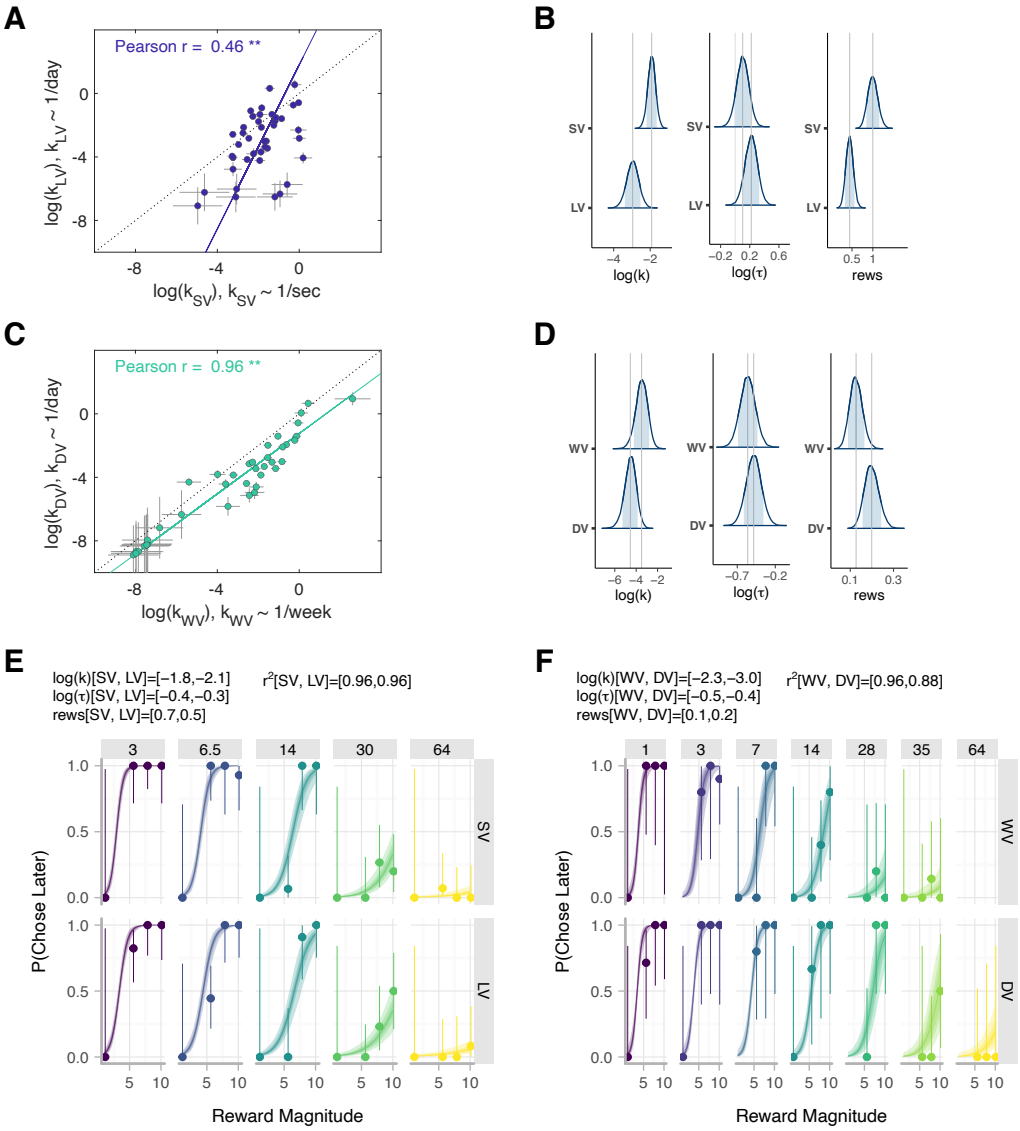

Figure 2. Comparison of discounting across tasks. (A)-(B) DS, sessions 1 & 2 with delays in days and seconds. (C)-(D) DW, session 3 with delays in days and weeks. (A)-(C) Each circle is one subject ($N = 41$). Discount factors are estimated in days but then converted to the units of the task. The error bars are the SD of the estimated coefficients. The solid line represents the perpendicular (or total) least squares (TLS) regression line (Huang et al., 2013). The dotted line is the unity line ($y = x$). (A) The logs of discount factors ($\log(k)$) in short delay verbal task (SV, x-axis) plotted against the logs of discount factors in long delay task (LV, y-axis). (B) Distribution of posterior parameter estimates of population-level $\log(k)$, decision noise, $\log(\tau)$, and reward scale parameter, rews, from the model fit for the SV and LV tasks. (C) The logs of discount factors in week delay verbal task (WV, x-axis) plotted against the logs of discount factors in day delay verbal task (DV, y-axis). (D) Distribution of posterior parameter estimates of population-level $\log(k)$, decision noise $\log(\tau)$ and reward scale parameter, rews, from the model fit for the WV and DV tasks. (E)-(F) An example softmax-hyperbolic fit of one subject. In each panel, the marker and error bar indicate the mean and binomial confidence intervals of the subject's choices for that offer. The smooth ribbon indicated the BHM model fits (at 50, 80, 99% credible intervals). At the top of the subject plot we indicate the mean estimates of $\log(k)$, $\log(\tau)$ and rews for each task for that subject. We also indicate the Bayesian r^2 for each task. (E) An example subject fit for DS, sessions 1 & 2 with delays in days and seconds. (F) An example subject fit for DW, session 3 with delays in days and weeks.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 2. Kfold model comparison

	ΔELPD	ΔSE
sessions 1 & 2		
reward scale	0	0
noise per unit	-439.8	60.2
reduced	-483.8	56.5
session 1		
reward scale	0	0
noise per unit	-145.5	35.4
reduced	-176.5	35.4
session 2		
reward scale	0	0
reduced	-277.5	40.5
noise per unit	-288.6	48.3
session 3		
reward scale	0	0
noise per unit	-18.5	21.4
reduced	-27.5	18.6

Note. 10-fold cross validation was done using the `kfold` function in the `brms` R package. The `loo_compare` function in the `brms` package arranges the models in order, from best to worst. The second column shows the difference between the expected log pointwise predictive density (ΔELPD, Vehtari et al., 2017) of the best model with the model on that row.

Evidence of context dependent temporal processing

The three experimental sessions under consideration were within-subject. This allowed us to compare the delay discounting in days between DS and DW sessions (LV vs. LV3). The discount factors ($\log(k)$) were highly correlated (Pearson $r = .74$, $p < .001$), but at the same time significantly different according to a permutation test ($M_{DS} = -3.06$ & $M_{DW} = -4.53$, $p < 0.001$). Or, if taken apart, in Table 1 the correlation between delay discounting in the long delay task was significantly higher between sessions 2 and 3, than between sessions 1 and 2 (Pearson $r = .87$ compared to $r = .52$, respectively; significant difference according to 'cocor' tests, Materials and Methods). This result can be attributed to the increase in stability of the time preferences over the gained experience with the task, rather than any dependence on the temporal context.

In order to additionally check for context dependent temporal processing, we took a further look at early trials. If choices are dependent on temporal context then it may take some number of trials in each task for a participant to stabilize his behavior. Accordingly, we tested for the existence and stability of this adaptation effect in early trials for both DW and DS sessions. As in Lukinova et al. (2019), we found that in DS there was a small adaptation effect: subjects were more likely to choose the later option in early trials (first four) of SV task in Figure 3A, than in the late trials. In DW, weeks were not interpreted differently than days (Figure 3B).

We also did fits jointly for all three delay sessions combined. The best model, according to the 10-fold cross validation criteria was also the 'reward scale' model (SI, Table S2). According to the model point estimates, the correlation size between $\log(k)$ in different units increased not significantly for SV vs. LV, but decreased significantly for LV vs. WV, according to 'cocor' tests, compared to when estimated separately (Figures 3D-E). The correlation between jointly and separately estimated $\log(k)$ in the long delay task was almost perfect and higher between joint and DS sessions estimation (Pearson $r = .96$, $p < 2.2e^{-16}$) than correlation between $\log(k_{LV})$ combined sessions and DW session (Pearson $r = .87$, $p < .001$). Using the joint fits the adaptation effect from early trials (after the change of temporal context) to late trials remained: participants interpreted seconds as smaller than days and weeks with this model fits (Figure 3C).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

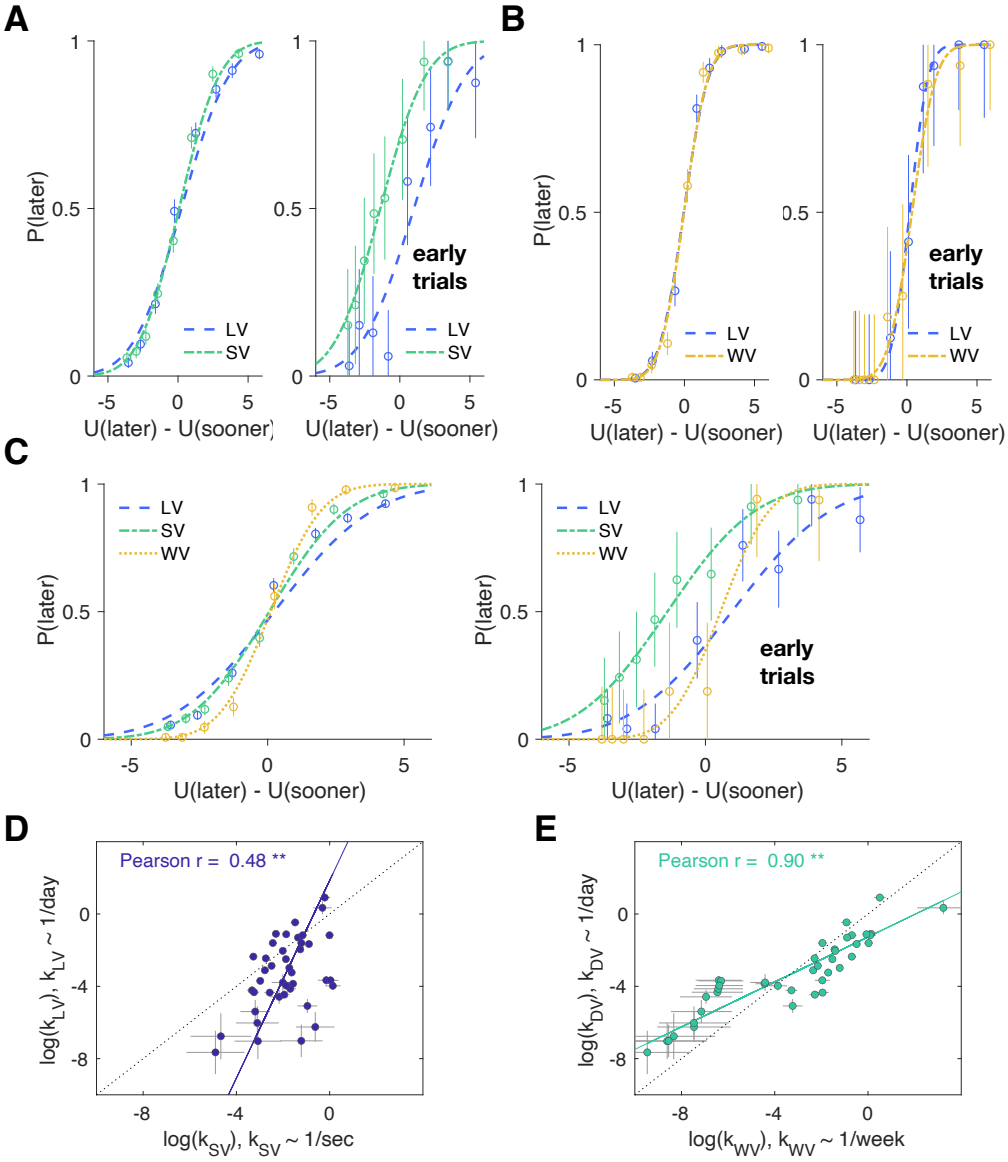

Figure 3. Early trials adaptation effect. (A)-(C) The offers for each subject were converted into a subjective utility, U, based on the subjects' discount factors in each task. This allowed us to combine data across subjects to plot psychometric curves of the probability of choosing the later option, P(later), for SV and LV averaged across all subjects comparing late trials (Trial in task > 5) (first facet) to the first four trials (early trials, second facet). (A) DS sessions fits; (B) DW session fits; (C) joint DS & DW session fits. (D)-(E) Sessions 1, 2 & 3 (DS & DW) joint fits with delays in days, seconds, and weeks. Each circle is one subject ($N = 41$). Discount factors are estimated in days but then converted to the units of the task. The error bars are the SD of the estimated coefficients. The solid line represents the perpendicular (or total) least squares (TLS) regression line. The dotted line is the unity line ($y = x$). (D) The logs of discount factors ($\log(k)$) in short delay task (SV, x-axis) plotted against the logs of discount factors in long delay task (LV, y-axis). (E) The logs of discount factors in week delay task (WV, x-axis) plotted against the logs of discount factors in day delay (long delay) task (DV = LV, y-axis).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Risk preferences are stable and consistent with the age group

The risk fits were done using the 'rho-beta' (standard power utility) and 'kappa-beta' (Balasubramani et al., 2014) Bayesian hierarchical models jointly for general and student (not considered in the paper) populations (Materials and Methods, Figure 4D&E). The comparison between the models was based on leave-one-out cross-validation criteria (using the `loo` function in the `loo` R package Vehtari et al., 2017) and favored the rho-beta model (ΔELPD = -578.7, ΔSE = 64.3 for the kappa-beta model compared to the rho-beta model). Given that ρ and κ (risk preference parameters) are strongly correlated (Pearson $r = -.95$, $p << 0.001$), we used the standard power utility model (rho-beta) point estimates further in Results. The subjects' choices were fit well by the rho-beta model as accessed using Bayesian r^2 and reported in each of the example subject risk plots in Figure 4A-C (see SI for all subject plots).

Subjects were consistent with the general notion of human subjects being risk averse and with previously reported risk preferences in the gains domain. Our subjects chose lottery 44.0% of the time and their average risk preference, $M_\rho = 0.75$ was slightly higher than expected (e.g., midlife adults choose lottery 42.9% of time; $M_\rho \approx 0.6$ for the same age group in Tymula et al., 2013).

Subjects' risk behavior was stable. We found no evidence that there was a difference in proportion of lottery choices between the two parts of the Risk task ($M_{r1} = 0.45$ & $M_{r2} = 0.44$, $p = 0.455$). Also we accessed subjects' rationality by presenting them with choices, in which one option was objectively better than the other. In these trials, subjects chose between a certain gain (surebet) of 4 coins and a lottery that offered a risky opportunity for gaining 4 or less coins. Since it is impossible to benefit by choosing the lottery, from economic standpoint the rational subjects must always choose the surebet regardless of their risk preferences. The frequency of first-order stochastic dominance violations (e.g., choosing any lottery with reward $<= 4$ coins) was negligible: 1.6% (which is much less than $\sim$5% reported for the same age group in Tymula et al., 2013) and not significantly different between two parts of the Risk task according to permutation test ($M_{r1} = 0.014$ & $M_{r2} = 0.016$, $p = 0.737$). Overall, 37% of subjects violated dominance at least once.

Subjects' risk preferences were not correlated with time preferences in seconds ($\log(k_{SV})$ vs. ρ, Pearson $r = -.02$, $p = 0.912$). Nevertheless, there was a small nonsignificant correlation of risk preferences with time preferences in days ($\log(k_{LV})$ vs. ρ, Pearson $r = -.29$, $p = 0.069$) and with difference between time preferences across time horizons ($\log(k_{SV}) - \log(k_{LV})$ vs. ρ, Pearson $r = -.31$, $p = 0.052$). Therefore, we checked whether there is any difference between subjects' risk preferences for whom $\log(k_{SV})$ (after converting units to seconds) was higher than $\log(k_{LV})$ compared to the opposite. Although the shift was visible in Figure 4F, the difference between risk preferences divided in that way was not significant (permutation test, $M_{\log(k_{SV})>\log(k_{LV})} = 0.795$ & $M_{\log(k_{SV})<\log(k_{LV})} = 0.642$, $p = 0.219$).

Stress measures convey distinct meanings and do not differ across genders

Apart from the experimental tasks, we also took several measures of stress from our participants. Importantly, we did not stress the subjects. Instead we collected the base level of stress using i) three questionnaires before the decision-making tasks during session 1 (PSS, BEPSI, LCU in Figure 5, Materials and Methods); ii) saliva samples before the decision-making tasks for each session that happened at approximately the same time and the same day of the week per participant (three samples per three sessions: log-transformed ls1-ls3 in Figure 5); iii) 'scarcity' questions at the start of decision-making task (ms and ts in Figure 5, Materials and Methods); and iv) hair samples two weeks after session 3 (three 1cm-segments analyzed separately: log-transformed lh1, lh2 and lh3 in Figure 5 from the stress level two months before collection to one-month-before stress and now, respectively). Due to voluntary nature of sample collection, we only had 34 subjects with all stress data.

Stress questionnaires can track 'stress' traits, saliva samples may provide a measurement of the cortisol concentration at a single point in time and cortisol of hair samples divided into 1cm length segments can correspond to chronic stress (one month of stress per one 1cm segment following Gow et al., 2010). According to the literature, we expected significant medium level correlation between hair and saliva measures around $r = 0.4$ (Vanaelst et al., 2012; van Holland et al., 2012; Zhang et al., 2018; Weckesser et al., 2019) and low level to no correlation between biological measures and questionnaires (Vanaelst et al., 2012; Stalder et al., 2017; Prado-Gascó et al., 2019). Some scholars suggest log-transformation of cortisol values to better approximate a Gaussian distribution of cortisol values across subjects (Lenow et al., 2017; Lempert et al., 2018a). Indeed, we

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

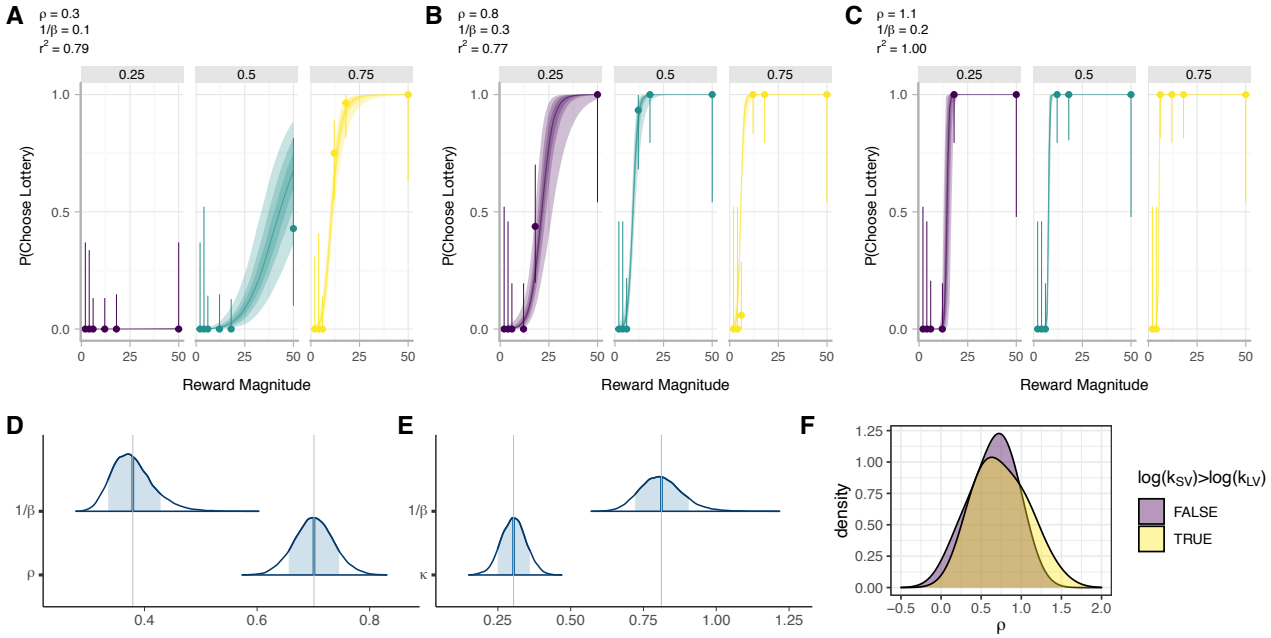

Figure 4. (A)-(C) An example rho-beta risk (power utility) fit of several subjects. In each panel, the marker and error bar indicate the mean and binomial confidence intervals of the subject's choices for that offer. The smooth ribbon indicated the BHM rho-beta model fits (at 50, 80, 99% credible intervals). At the top of the subject plot we indicate the mean estimates of risk preference parameter, ρ, and decision noise, $1/\beta$. We also indicate the Bayesian r^2. (A) A highly risk-averse subject; (B) a slightly risk-averse subject; and (C) a risk-seeking subject. (D) Distribution of posterior parameter estimates of the population-level risk preference parameter, ρ and decision noise, $1/\beta$, from the rho-beta risk model fit for the risk task. (E) Distribution of posterior parameter estimates of the population-level risk preference parameter, κ and decision noise, $1/\beta$, from the kappa-beta risk model fit for the risk task. (F) Density estimation of risk parameter, ρ, split by positive or negative difference between discount factors in short and long delay task.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

decided it is reasonable to modify all of our cortisol values based on the distributions before and after log-transformation (SI, Figure S5). From the correlation analysis where we combined all stress measures, cortisol levels were not associated with questionnaire scores, but questionnaire scores were associated with the 'scarcity' average self-reported ratings. Overall, we found more significant correlations within the stress measure groups (which may validate stress measures per se), compared to across stress measure groups (Figure 5A). Still, saliva cortisol averaged across three sessions (ls123) and from the third session (in Figure 5A) was positively correlated with two hair cortisol samples as literature suggests (ls123 vs. lh2: Pearson $r = 0.38$, $p = 0.024$; ls123 vs. lh3: Pearson $r = 0.35$, $p = 0.039$). With the multiple comparison correction (10 comparisons per stress variable) only positive correlations between lh2 and lh3 (Pearson $r = 0.46$, $p = 0.005$); PSS and BEPSI questionnaires (Pearson $r = 0.78$, $p < 0.001$); and BEPSI and ms (Pearson $r = 0.52$, $p < 0.001$) remained significant.

We also ran two principal component analyses (PCA). First, we used PCA to visualize the relationships across the stress groups. To this end, we used the eleven stress-vectors (three saliva, three hair, three questionnaire data points and two 'scarcity' ratings averages) for each subject in Figure 5B-C. Above 65% of variance was explained by the first four principle components and the first 4 components had an eigenvalue > 1 (Figure 5C). The second principle component along the eleven stress-vectors distinguished between human bio samples (hair and saliva) and questionnaires, whereas the third principle component distinguished between hair and saliva (Figure 5B). According to the angles (calculated from the positive x-axis following the four quadrants of 2D plane) in Table 3 we got some evidence that clusters of stress groups were different. However, we could not say the same for the gender groups: no clustering was visible. Literature suggests that because gender is associated with neuroendocrinological substrates (e.g., cortisol, Takahashi, 2004; Barel et al., 2017), the relationships between stress hormones and decision making should control for gender. Moreover, there are studies that only include males due to reported menstrual and contraceptive effects on cortisol levels (Lempert et al., 2012; Stalder et al., 2017). First, we visually assessed the differences between samples (by gender), given PC1-PC2 and PC2-PC3 planes and found that the gender ellipses as well as the data are highly overlapping. Second, we used permutation tests to compare stress measures between gender categories. We did not find effects of gender for most of the stress variables ($p > .1$). The stress measures were significantly different between genders only for saliva cortisol for session 1 (s1, $M_f = 0.89$ & $M_m = 1.51$, $p = 0.013$).

Second, PCA was used to reduce the dimensionality of the stress data and its changes (including fluctuations) while retaining most of the variation in the data. To this end, we used the 23 stress-vectors (nine saliva: three cortisol levels log-transformed, three deltas, and three fluctuations; nine hair; three questionnaire data points; and two 'scarcity' ratings averages) for each subject. The first eight components had an eigenvalue > 1 and explained more than 82% of variance. However, even the first four components explained more than 56% of variance. We used the first four principal components from this analysis for results further.

Table 3. PCA angles

	PC1-PC2 plane	PC2-PC3 plane
ls1	207°	148°
ls2	239°	141°
ls3	244°	160°
lh1	180°	270°
lh2	223°	209°
lh3	247°	188°
pss	143°	13°
bepsi	164°	27°
lcu	137°	293°
ms	161°	47°
ts	127°	24°

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

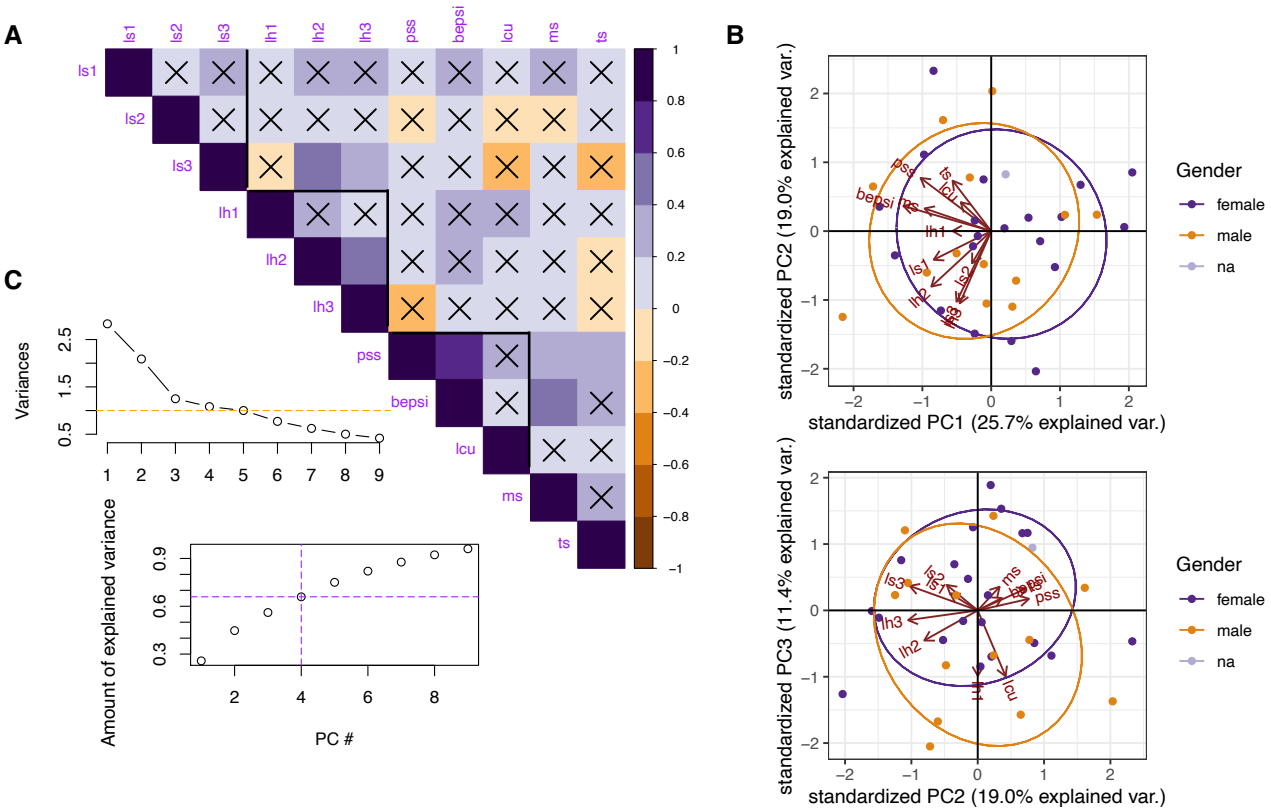

Figure 5. N = 34 (A) Correlation plot across all stress measures. The color intensity (vertical colorbar) signifies the correlation r coefficient. The insignificant correlations, e.g. where $p >= 0.05$ are indicated with an 'x'. The black borders are meant to highlight different types of stress measures and correlations within each stress measure. (B) Two 2D planes with axes being three first principle components (PCs) (Top: PC1-PC2 plane; Bottom: PC2-PC3 plane). These are PCA biplots that combine standardized data and the stress variables loadings (dark red arrows, with directions of corresponding eigenvectors). The data color is based on the gender group. The ellipses surrounding the data points (with color corresponding to the gender group) are drawn around clusters identified after automatic classification, based on PCA scores. (C) Top: Screeplot of the Eigenvalues of all PCs. The eigenvalue (y-axis) of each PC (x-axis - PC number). The cutoff is drawn at eigenvalue that equals 1. Bottom: Cumulative variance plot. The total variance explained (y-axis) depending on how many PCs to account for. The cutoff is drawn as dashed vertical line at 4 PCs.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Only hair cortisol is associated with model-based measures

We did not find significant correlations between model-based measures of economic preferences and saliva cortisol, questionnaire scores, or scarcity ratings. The only significant associations we found were between short delay discount factor and change (including absolute change) in cortisol level from hair sample (e.g., $\log(k_{SV})$ vs. Δlh32 Pearson $r = -0.41$, $p = 0.012$); between decision noise in the risk task and change (including absolute change) in cortisol level from hair sample (e.g., $1/\beta$ vs. Δlh32 Pearson $r = -0.50$, $p = 0.002$); and between risk aversion parameter and fluctuation (i.e. absolute change) in cortisol level from hair sample (ρ vs. flh32 Pearson $r = 0.48$, $p = 0.003$). Also, $\log(k_{SV})$ was negatively correlated with PC2 (one of the PCs following dimensionality reduction from 23 to 8 across all stress measures and their changes). All significant correlations are highlighted in bold in Table 4. Still, most of the significant correlations (especially in the risk task) were driven by a single outlier in the risk task, both risk-seeking and decision noise for whom were extreme (SI, Figure S7). The only significant relationship that remained was a positive correlation between stress fluctuations using hair sample and delay discounting in the short delay task (Pearson $r = 0.40$,$p = 0.020$.

Next, we looked at associations between the change in discount factors for short and long delay tasks across sessions and stress measures. We found significant positive correlations between the change in the short delay discount factor ($\log(k_{SV2}) - \log(k_{SV1})$) and saliva cortisol for the second session (ls2), hair cortisol change (Δlh3), and time scarcity (ts). Also, the change in the long delay discount factor ($\log(k_{LV2}) - \log(k_{LV1})$) was positively correlated with time scarcity. Nevertheless, these correlations were moderate (Pearson $r \sim 0.34$, all $p > 0.01$) and would no longer be significant if we used the multiple comparison correction.

The model-free short delay task measures are moderately related to stress measures

Previous studies have shown that the magnitude of the physiological response to stress can predict behavior in PFC-dependent tasks, importantly, with both linear (Goldfarb et al., 2017; Maier et al., 2015) and curvilinear (Luksys and Sandi, 2011; Lempert et al., 2018a) relationships being possible. To that end, we explored several other per task / session measures, including subjects' profit, total waiting time (wait), proportion of later choices (PL, or lottery choices for risk task) and reaction time (RT) in relation to stress levels and their change. It is important to note that the total waiting time and the profit variables were experienced only in the short verbal task (SV) and in the other tasks only represent potential outcomes (hypothetical total) compared to only one random outcome selected for payout.

Across these behavioral measures (except the RTs) there were strong positive significant relationships (as measured by correlation analysis, Figure 6). This was expected from the delay task design, since the more often a subject chose a later option, the more she would earn and the more she would have to wait (if all trials are counted towards the total, e.g., between profitS and waitS, Pearson $r = 0.99$, $p < 2.2e^{-16}$). Interestingly, RTs in SV were positively correlated with other model-free measures: the more subject spent time on decision the more he earned or chose later option in the session 1 (e.g., RT_S1 vs. profitS1, Pearson $r = 0.38$, $p = 0.015$). In other tasks the relationship between RT and other behavioral variables was not observed.

Between these model-free task measures and stress variables significant correlations were observed only for the short delay task compared to other tasks, with the only exception of proportion of lottery choices in the risk task being positively correlated with the fluctuation of cortisol across hair samples (PL_R vs. lfh32: Pearson $r = 0.36$, $p = 0.032$). For example, the proportion of later choices in SV (PL_S) was negatively associated with LCU (Pearson $r = -0.34$, $p = 0.038$), but positively associated with changes in hair cortisol level (Δlh32, Pearson $r = 0.38$, $p = 0.023$) and the second principle component (PC2, Pearson $r = 0.42$, $p = 0.015$). In addition, we tested the changes of model-free task measures (i.e. PL) over time against stress deltas. We found just a few significant relationships: changes in hair cortisol were strongly and negatively associated with changes in choosing later option in SV (ΔPL_S2 = PL_S2 - PL_S1 vs. Δlh2 Pearson $r = -0.34$, $p = 0.046$; ΔPL_S2 vs. Δlh3 Pearson $r = -0.49$, $p = 0.003$). Nevertheless, the significance of all of these relationships would not hold if one used a multiple comparison correction (e.g., Bonferroni $p = 0.05/27 = 0.002$ for 27 stress variables being compared at each time).

We did not stress our subjects and only looked at changes in cortisol levels compared to baseline session or month. Still, following Lempert et al. (2018a) we checked whether we can find any curvilinear relationships for

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 4. Correlations of subjects' decision-making parameters with stress variables

	$\log(k_{SV})$	$\log(k_{LV})$	$\log(\tau_{SV})$	ρ	$1/\beta$
Saliva cortisol: $N = 41$					
log(s1)	$r = -0.16$	$r = 0.02$	$r = -0.06$	$r = 0.15$	$r = 0.02$
	$p = 0.310$	$p = 0.905$	$p = 0.721$	$p = 0.347$	$p = 0.902$
log(s2)	$r = -0.10$	$r = -0.03$	$r = -0.05$	$r = 0.23$	$r = 0.20$
	$p = 0.527$	$p = 0.869$	$p = 0.733$	$p = 0.157$	$p = 0.203$
log(s3)	$r = -0.07$	$r = 0.15$	$r = 0.05$	$r = 0.05$	$r = -0.15$
	$p = 0.671$	$p = 0.364$	$p = 0.776$	$p = 0.747$	$p = 0.360$
Δls2 = log(s2)−log(s1)	$r = 0.07$	$r = -0.03$	$r = 0.01$	$r = 0.02$	$r = 0.11$
	$p = 0.658$	$p = 0.840$	$p = 0.934$	$p = 0.925$	$p = 0.494$
Δls3 = log(s3)−log(s1)	$r = 0.09$	$r = 0.10$	$r = 0.09$	$r = -0.09$	$r = -0.13$
	$p = 0.564$	$p = 0.552$	$p = 0.588$	$p = 0.556$	$p = 0.412$
Δls32 = log(s3)−log(s2)	$r = 0.02$	$r = 0.14$	$r = 0.08$	$r = -0.12$	$r = -0.27$
	$p = 0.918$	$p = 0.381$	$p = 0.627$	$p = 0.457$	$p = 0.090$
fls2 = abs(log(s2)−log(s1))	$r = -0.17$	$r = -0.04$	$r = 0.11$	$r = 0.03$	$r = 0.05$
	$p = 0.293$	$p = 0.824$	$p = 0.504$	$p = 0.875$	$p = 0.780$
fls3 = abs(log(s3)−log(s1))	$r = 0.02$	$r = 0.02$	$r = 0.13$	$r = 0.12$	$r = 0.09$
	$p = 0.899$	$p = 0.900$	$p = 0.411$	$p = 0.467$	$p = 0.578$
fls32 = abs(log(s3)−log(s2))	$r = -0.11$	$r = -0.18$	$r = 0.23$	$r = -0.06$	$r = 0.07$
	$p = 0.480$	$p = 0.260$	$p = 0.156$	$p = 0.701$	$p = 0.650$
Hair cortisol: $N = 36$					
log(h1)	$r = 0.23$	$r = -0.02$	$r = 0.16$	$r = 0.28$	$r = 0.27$
	$p = 0.183$	$p = 0.890$	$p = 0.366$	$p = 0.097$	$p = 0.105$
log(h2)	$r = 0.25$	$r = 0.20$	$r = 0.05$	$r = 0.28$	$r =$ **0.34**
	$p = 0.136$	$p = 0.239$	$p = 0.769$	$p = 0.102$	$p =$ **0.043**
log(h3)	$r = -0.17$	$r = 0.13$	$r = -0.27$	$r = 0.02$	$r = -0.17$
	$p = 0.311$	$p = 0.458$	$p = 0.117$	$p = 0.909$	$p = 0.323$
Δlh2 = log(h2)−log(h1)	$r = -0.01$	$r = 0.17$	$r = -0.10$	$r = -0.04$	$r = 0.01$
	$p = 0.941$	$p = 0.327$	$p = 0.570$	$p = 0.807$	$p = 0.958$
Δlh3 = log(h3)−log(h1)	$r = -0.29$	$r = 0.10$	$r = -0.29$	$r = -0.21$	$r = -0.32$
	$p = 0.088$	$p = 0.570$	$p = 0.088$	$p = 0.213$	$p = 0.054$
Δlh32 = log(h3)−log(h2)	$r =$ **-0.41**	$r = -0.08$	$r = -0.30$	$r = -0.26$	$r =$ **-0.50**
	$p =$ **0.012**	$p = 0.624$	$p = 0.080$	$p = 0.127$	$p =$ **0.002**
flh2 = abs(log(h2)−log(h1))	$r = 0.23$	$r = -0.01$	$r = -0.08$	$r = -0.23$	$r = -0.09$
	$p = 0.177$	$p = 0.952$	$p = 0.625$	$p = 0.174$	$p = 0.614$
flh3 = abs(log(h3)−log(h1))	$r =$ **0.50**	$r = 0.03$	$r = 0.31$	$r = 0.03$	$r = 0.24$
	$p =$ **0.002**	$p = 0.848$	$p = 0.065$	$p = 0.847$	$p = 0.157$
flh32 = abs(log(h3)−log(h2))	$r = 0.23$	$r = -0.23$	$r = 0.20$	$r =$ **0.48**	$r =$ **0.58**
	$p = 0.170$	$p = 0.180$	$p = 0.247$	$p =$ **0.003**	$p =$ **0.0002**
Stress questionnaires: $N = 38$					
PSS	$r = 0.15$	$r = -0.13$	$r = -0.05$	$r = 0.05$	$r = 0.01$
	$p = 0.384$	$p = 0.430$	$p = 0.760$	$p = 0.762$	$p = 0.942$
BEPSI	$r = 0.20$	$r = -0.03$	$r = -0.10$	$r = 0.11$	$r = 0.02$
	$p = 0.231$	$p = 0.856$	$p = 0.534$	$p = 0.511$	$p = 0.917$
LCU	$r = 0.31$	$r = 0.02$	$r = -0.07$	$r = 0.04$	$r = 0.04$
	$p = 0.056$	$p = 0.893$	$p = 0.683$	$p = 0.820$	$p = 0.832$
money scarcity ($N = 40$)	$r = -0.08$	$r = -0.19$	$r = -0.03$	$r = 0.26$	$r = 0.12$
	$p = 0.603$	$p = 0.233$	$p = 0.870$	$p = 0.102$	$p = 0.452$
time scarcity ($N = 40$)	$r = 0.05$	$r = 0.01$	$r = -0.22$	$r = 0.21$	$r = 0.15$
	$p = 0.746$	$p = 0.971$	$p = 0.182$	$p = 0.191$	$p = 0.355$
PCs: $N = 34$					
PC1	$r = 0.21$	$r = -0.15$	$r = 0.20$	$r = 0.04$	$r = 0.26$
	$p = 0.232$	$p = 0.396$	$p = 0.268$	$p = 0.843$	$p = 0.144$
PC2	$r =$ **-0.43**	$r = -0.04$	$r = -0.10$	$r = -0.34$	$r = -0.33$
	$p =$ **0.011**	$p = 0.833$	$p = 0.577$	$p = 0.051$	$p = 0.056$
PC3	$r = -0.22$	$r = -0.15$	$r = -0.33$	$r = -0.05$	$r = -0.10$
	$p = 0.206$	$p = 0.405$	$p = 0.060$	$p = 0.784$	$p = 0.567$
PC4	$r = 0.06$	$r = -0.33$	$r = -0.06$	$r = -0.02$	$r = -0.02$
	$p = 0.724$	$p = 0.055$	$p = 0.719$	$p = 0.904$	$p = 0.922$

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

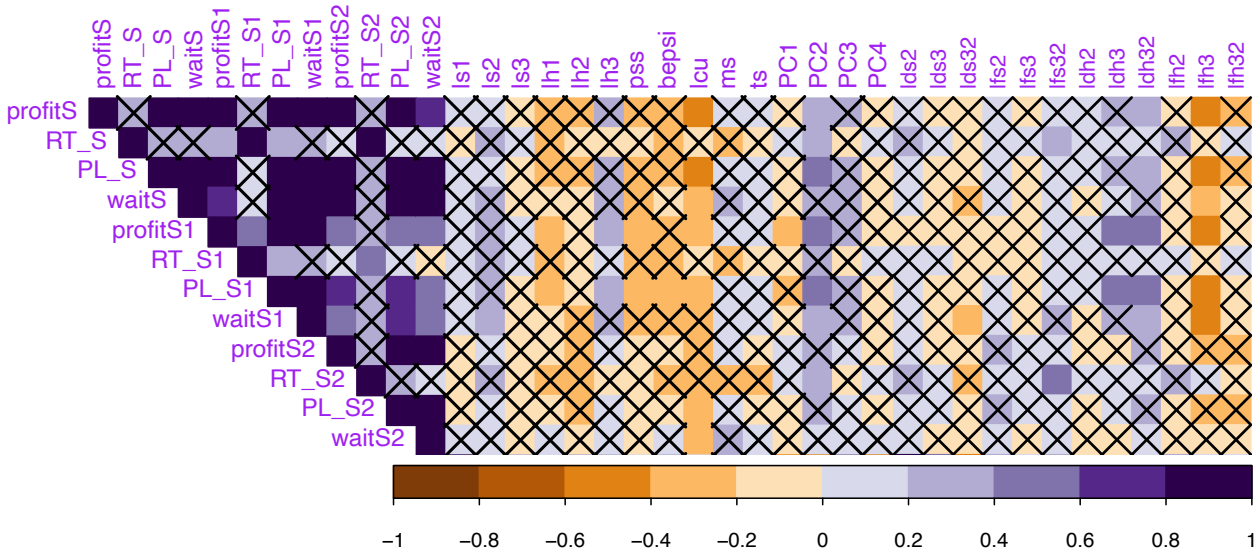

Figure 6. Correlation analysis across model-free measures in short verbal task and stress measures. The color intensity (horizontal colorbar) signifies the magnitude of the correlation. If the cell does not have a cross on it, this correlation is significant at .05 level.

those with increase in biological stress level compared to the baseline. There were only 13 subjects with such positive changes from lh1 to lh2 (Δlh2), 16 subjects from lh1 to lh3 (Δlh3) and 17 subjects from lh2 to lh3 (Δlh32). Out of many possible comparisons we found just a few inverted-U-shaped relationships between hair stress deltas and model-free parameters. Out of six potential relationships selected in the exploratory analysis (SI, Figure S6), only one held after likelihood ratio tests (change in likelihood was significant using a χ^2 statistic). The complex model (polynomial of degree 2) was preferred to the nested one (behavior $\sim \Delta$stress) and, therefore the curvilinear was preferred to the linear relationship for:

RT_S $\sim \Delta lh2 + (\Delta lh2)^2$, $P(> \chi^2) = 0.054$, $\beta(\Delta lh2)^2) = -0.62$, $p = 0.100$.

Discussion

We confirmed that in the general population discount factors across time-horizons were highly related as in the students population in our studies before (Lukinova et al., 2019; Lukinova and Erlich, 2021). In this paper, however, we improved our modeling. The best hierarchical Bayesian model was the one that allowed re-scaling reward and at the same controlling for the units of the task in both discount factor and decision noise. Also, this time (compared to Lukinova et al., 2019) all tasks were done within subject, so we were able to compare behaviors in the long delay task (delays in days) when the other task was in seconds (sessions 1 & 2) and in weeks (session 3). We observed a temporal effect that makes the delays in days in the long delay task paired with the delay task in seconds seem longer than the delays in weeks in Figure 3C.

Many scholars point to benefits of accounting for risk preferences in the delay discounting (Lempert et al., 2018c). Here, risk task was also included in the battery of decision-making tasks. We found risk preferences to be more correlated with the time preferences from the long delay task, but not short delay task. Once we used power utility model inside the delay discounting model in order to de-correlate the risk preferences and time preferences in days (Lopez-Guzman et al., 2019), the opposite relationship appeared: risk parameter ρ was correlated with the discount factor in the short delay task and not correlated with that in the long delay task (SI). Therefore, we suggest that a more complex model of joint estimation is required, since risk preferences should be accounted for in 'postponing' but not in 'waiting' tasks. We found some evidence towards the 'waiting/postponing story': those who had smaller discount factor in the long delay task than predicted through the discount factor in the short delay task were more risk-seeking (Figure 4F). This was indeed, embedded in our design, since the short delay task did not include risk: participant received the sum of all

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

coins she earned from all trials and had already waited for all delays in seconds to get those coins. On the contrary, in the long delay task, one random choice was chosen and by choosing the later option participant would risk to be paid at a much later date: from several days to several months.

In our study, biological stress measures were moderately correlated between each other as seen in the literature before, further validating the stress measurement system (Vanaelst et al., 2012; van Holland et al., 2012; Zhang et al., 2018; Weckesser et al., 2019). However, overall we found more significant relationships within stress groups, rather than across them. Also, following our principle component analysis, distinct stress measures were well clustered on the 2D planes of the first three principle components. This gives more evidence that self-reported stress, saliva cortisol and hair cortisol might affect decision making and operate via distinct channels.

We found only nominal relationships of model-based parameters with stress. In particular, we found that stress measured via a hair sample was more related to the estimated parameters, especially for the short delay task: e.g., the higher was the fluctuation, i.e. the absolute change, in the stress level the steeper was the discounting in the short delay task, whereas significant relationships between stress and the risk task estimates were driven by an outlier.

Previous literature accepts both linear and curvilinear relationships between changes in cortisol as a physiological response to stress and behavior (Goldfarb et al., 2017; Maier et al., 2015). Model-free task measures constructed for testing such relationships, indeed, showed significant associations with all stress variables, including self-reported stress via questionnaires. Importantly, most of the significant results from this analysis were between model-free measures in the short delay task and stress variables. This is the only task where all trials contributed to the total profit (and waiting time, for delay tasks). In other words, baseline stress level and its fluctuations were more prone to influence behavior in 'real' experiential task rather than in a task with potential outcomes determined by a random draw in the end of the experimental session. We observed that the smaller were changes in hair cortisol level between the months of the study and higher were the life change units (from the questionnaire) the less subjects were choosing the later option. Among curvilinear relationships we did not confirm previous results about subjects' earnings (Lempert et al., 2018a), but found some evidence for the reaction times: when chronic stress level increased slightly, it took longer for participants to decide in the short delay task, but when the cortisol measure from hair samples increased to higher levels, they decided quicker whether to pick a sooner or later option. Nevertheless, tests for curvilinear relationships included only a small portion of participants, those for whom there was an increase in biological stress level compared to the baseline.

The longitudinal within-subject design allowed us to compare changes in biological stress markers over time to changes in behavior. We did not find evidence that changes in model-based parameters for the intertemporal choice task were associated with the changes and fluctuations of stress from the baseline. Nor did we find any stress-recency effects using model-free task measures by creating the model-free task deltas for the short delay task.

In our study we did not stress our participants. Instead, we were looking at cortisol levels, its changes and fluctuations from the baseline that may be due to individual responses to the environment. This is rather new. Previously in control groups random fluctuations of cortisol did not yield the same behavioral patterns (Lempert et al., 2018a). However, most of the previous studies were using saliva measure (Takahashi, 2004; Linz et al., 2018). Most of our significant relationships between stress and decision making included hair cortisol measure. We conclude that for subtle changes in behavior due to random cortisol fluctuations hair stress measure might be more reliable.

Hormones are tricky to measure and administer. They change across the day and across the month. For example, stressors that are uncontrollable tend to produce a high overall level of daily cortisol release, whereas controllable stressors - higher-than-normal morning cortisol (Miller et al., 2007). As a result, many studies involving hormones and decision-making produce conflicting results or fail to be replicated and, therefore, are not able to advance the current state of the discipline. Our results did not support the hypotheses that increased stress leads to steeper discounting and increased risk aversion. Instead, we found that endogenous subclinical variation in cortisol only weakly correlates with economic preferences. Moreover, the strongest link we found was that fluctuations in cortisol levels correlated with steeper discounting in the short delay task.

Additional multi-task and multi-stress studies with higher N are required to establish effects discovered in

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

this paper. An accurate understanding of the biological processes that lead people to be more risk averse and more likely to discount future payoffs is still tenuous at best. Unraveling mechanisms of how chronic stress spreads unfavorable effects to behavior is of critical importance to policymakers and of strong interest to the general public.

Materials and Methods

Participants

In this study, we recruited 44 participants total (26 female, 17 male, 1 other) including 15 participants from NYU Shanghai staff and 29 participants from Chinese working population around New York University (NYU) Shanghai and East China Normal University (ECNU) campuses. We pre-screened subjects before recruitment and excluded smokers, drinkers, and subjects taking medicine, or suffering from acute or chronic hormonal dysregulations. Three participants withdrew after the first session and the rest (41: 23 female, 17 male, 1 other, $M_{age} = 30$) completed the whole experiment. All of our participants shared a similar socioeconomic status, but had diverse consumption level (SI, Figure S1A). 31 participants (out of 41) had no siblings, 22 (out of 41) were not married and 32 (out of 41) did not have any chronic illnesses. The data collection for the study was done in spring and summer of 2018. The study was approved by the IRB of NYU Shanghai. All subjects gave informed consent before participation in the study.

Experimental Tasks

There were three behavioral task sessions in total. They were scheduled approximately every two weeks and took place in the NYU Shanghai Behavioral and Experimental Economics Laboratory, or the experimental room in the Geo Building at ECNU. Each session consisted of a two-alternative choice task. In the first two sessions, participants participated only in the intertemporal choice task and made their choices between a later option (delay in seconds and days) and an immediate option (now and today). The last session involved the intertermporal choice task comprising choices between a later option (delay in days and weeks) and an immediate option (today) and the risk task, where choice was made between a risky option (a probabilistic reward) and a safe option (surebet). Participants received a 30 CNY ($4.3 USD) per session participation fee as well as up to an additional 70 CNY ($9.9 USD) per session based on their individual performance in the tasks.

Stress

We used three questionnaires to measure perceived stress. The 10-item version of the Perceived Stress Scale (PSS, Cohen et al., 1983) was used to assess an individual's subjective appraisal of particular life events/situations as being unpredictable, uncontrollable, and/or overloaded. On a 5-point Likert scale ranging from 0 'never' to 4 'very often', participants rated how often in the previous month they felt or thought as described in 10 examples. The Brief Encounter Psychosocial Instrument (BEPSI, Frank and Zyzanski, 1988) usually consists of 6 questions. We omitted the first open-ended question and used the rest: 5 closed-ended questions. Subjects were asked to respond yes or no to each item and if yes, to rate the impact of these stressors on a scale of 1 to 10. The Social Readjustment Rating Scale (SRRS, Holmes and Rahe, 1967) was used for identifying major stressful life events. Each one of the 43 stressful life events was awarded a Life Change Unit (LCU) depending on how traumatic it was felt to be by a large sample of participants. The total then was calculated by adding all LCUs for the events a subject experienced. Thus, in all questionnaires higher scores were associated with higher levels of perceived stress. All questionnaires were administered in Chinese.

We also asked three 'scarcity' questions in Chinese in the beginning of each session since lacking money or time resources could also lead to stress (Mullainathan and Shafir, 2013):

- To what extent did you feel that you were out of money in the last two weeks?
- To what extent did you feel that you were out of time in the last two weeks?
- To what extent are you in a hurry today?

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Subjects answered the above questions using a 7-point rating scale (from lowest = 1 to highest = 7). The answers for the same question across sessions as well as between the last two questions were highly correlated (all $p < 0.005$), so we decided to keep two averages across three sessions for the money and time scarcity (ms and ts, respectively).

In order to obtain another measure of participant's stress level, in the beginning of each session, we collected a saliva sample from each participant. For each participant, the collection time remained the same during three sessions (as in Pruessner et al., 2003; González-Cabrera et al., 2014). Therefore, we could compare saliva cortisol levels to the baseline (session 1) and attribute these Δs to the changes from the base cortisol level. Saliva samples were collected using the Salivette systems. The Salivette systems (Sarstedt, Shanghai) came with synthetic swabs. Subjects were asked to refrain from eating, drinking and/or brushing teeth for at least 1 hour before sampling and to rinse their mouth 10 minutes before arriving to the experimental facility. Before each experimental session, experimenter asked a subject to place the swab in the mouth, chew it for 1 min and then transfer it into the tube. All subjects had three saliva samples of necessary quality. The samples were stored on ice till the end of the experimental day, then, were centrifuged and stored at -20 °C until analysis at NYU Shanghai Molecular Biology Lab. On the day of analysis, we measured salivary cortisol by enzyme-linked immunosorbent assay (ELISA, IBL International, Germany). Saliva samples were assayed in duplicate (40 * 2 samples per ELISA kit) and measured according to the instructions of the kits. The controls were within expected bounds.

Two weeks after all three experimental sessions were completed by subjects, we also collected their hair samples to measure the levels of chronic stress. 36 subjects returned to provide hair samples. Hair samples were collected by cutting two strands of hair (with the total thickness of a toothpick) from posterior vertex area. All samples were stored in aluminium foil at room temperature. LC-MS/MS atmospheric pressure chemical ionization (ACPI, MRM mode) was used to analyse cortisol in hair samples. The analysis was performed by Dr. Huihua Deng's lab at Southeast University using their standard approach (Chen et al., 2013). Hair strands longer than 3 cm were cut as close to the scalp as possible and segmented into three 1cm-long parts. All segmented samples were labelled as SXxx-01- SXxx-03 from bottom (closest to the scalp) to top. Totally 36*3 = 108 samples were analysed. Recovery rate was within 90-105%, and intra-day and inter-day coefficients of variation were below 10%, which satisfied measurement requirement.

Analysis

The discount factors were estimated using a softmax-hyperbolic fit (i) in a similar way as in Lukinova et al. (2019), i.e. a four population level and three subject level parameters model - reduced model; (ii) with a noise per unit model; and (iii) with a reward scale model, where an additional parameter 'rews' scaled the delayed reward per unit. All mixed-effects models were done after conversion of all delays to days (resulted in the same fits as fits in the units of the task) with the 'brms' package in R (Bürkner, 2017) that allows to do BHM of nonlinear multilevel models in Stan (Carpenter et al., 2016) with the standard R formula syntax.

The reduced model ($M_{4p,3s}$) had 4 population level parameters ($\log(k)$, and decision-noise, $\log(\tau)$ for each of the two intertemporal choice tasks) and 3 parameters per subject: $\log(k_{SV})$, $\log(k_{LV})$ and $\log(\tau)$.

```
n_chose_later | trials(n_trials) ~
(later_reward/(1 + exp(logk)*delay)-sooner_reward)/exp(logtau),
logtau ~ unit + (1 | subjid),
logk ~ unit + (unit | subjid)
```

The noise per unit model ($M_{4p,4s}$) had 4 population level parameters ($\log(k)$, and decision-noise, $\log(\tau)$ for each of the two intertemporal choice tasks) and 4 parameters per subject: $\log(k_{SV})$, $\log(k_{LV})$, $\log(\tau_{SV})$ and $\log(\tau_{LV})$.

```
n_chose_later | trials(n_trials) ~
(later_reward/(1 + exp(logk)*delay)-sooner_reward)/exp(logtau),
logtau ~ unit + (unit | subjid),
logk ~ unit + (unit | subjid)
```

The reward scale model ($M_{6p,3s}$) had 6 population level parameters ($\log(k)$, reward scale parameter, rews, and decision-noise, $\log(\tau)$ for each of the two intertemporal choice tasks) and 3 parameters per subject: $\log(k_{SV})$, $\log(k_{LV})$ and $\log(\tau)$.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

```
n_chose_later | trials(n_trials) ~ (exp(rews)*later_reward / (1 + exp(logk) * delay) - sooner
    _reward) / exp(logtau),
logtau ~ unit + (1|subjid),
logk ~ unit + (unit|subjid),
rews ~ unit
```

where `later_reward` is the later reward, `sooner_reward` is the sooner reward; `logk` is the natural logarithm of the discounting parameter k and `logtau` ($\log(\tau)$) is the natural logarithm of the decision noise; in the binomial specification (the bernoulli specification resulted in the same fits, SI) the data was grouped and summarized by distinct trial types, where `n_chose_later` is the count of choices when the later reward is selected and `n_trials` is the number of trials for a particular trial type. All model syntaxes above were valid for fitting two sessions or one session that include two delay-discounting tasks (three session fits are detailed in SI), where SV and LV will be changed to DV(LV) and WV for session 3.

The risk preferences were estimated using a rho-beta (power utility) and kappa-beta risk models. The models ($M_{2p,2s}$) had 2 population level parameters: ρ (or κ), risk preference parameter, and decision-noise, $1/\beta$ and the same 2 parameters per subject. All mixed-effects models are done with the 'brms' package in R with the standard R formula syntax:

```
n_chose_later | trials(n_trials) ~ beta*(probability*(lottery_mag)^r - (sb_mag)^r),
beta ~ (1|subjid),
r ~ (1|subjid)

n_chose_lot | trials(n_trials) ~ beta*(probability*(lottery_mag)-kappa*sqrt(lott_var) - sb_
    mag),
beta ~ (1|subjid),
kappa ~ (1|subjid)
```

where `lottery_mag` is the probabilistic reward, `sb_mag` is the surebet; `r` is ρ (or `kappa` is κ) - the risk preference parameter, with ρ values higher than 1 (or $\kappa < 0$) values indicative of increased risk tolerance and less than one of risk aversion (or $\kappa > 0$, which strictly speaking is indicative of variance aversion) and inverse `beta` is the decision noise; `lott_var`, or lottery variance was calculated asymptotically as

```
lott_var = probability * (1 - probability) * (lottery_mag)^2
```

As before for delay-discounting models, for the binomial specification the data was grouped and summarized by distinct trial types, where `n_chose_later` is the count of choices when the later reward is selected and `n_trials` is the number of trials for a particular trial type. One disadvantage of the rho-beta model is stronger correlation between ρ and respective β (Pearson $r = -.42$, $p = 0.007$) compared to that for the kappa-beta model (Pearson $r = -.22$, $p = 0.159$).

The significant difference in correlations was tested using R package 'cocor' (Diedenhofen and Musch, 2015) assuming overlapping (correlations that have one variable in common) or nonoverlapping dependent correlations. The visualizations of the correlation matrices were done via R package 'corrplot' (Friendly, 2002). The principle component analysis was performed with R function 'prcomp' from package 'stats'. The regression analysis and tests of curvilinear relationships used R package 'lme4' (Bates et al., 2015). The permutation tests of differences between the means of two groups were done by shuffling the group label and computing the mean between the shuffled groups 10000 times. This generated a null distribution which was used to estimate the probability of observing the true difference between groups (bootmean in https://github.com/erlichlab/elutils).

Software

The code for the delay-discounting and risk tasks was written in Python using the 'PsychoPy' toolbox (version 1.83.04; Peirce, 2007) available at https://www.github.com/erlichlab/gpstress/src/task. All analyses and statistics were performed either in Matlab (version 9.3, or higher, The Mathworks, MA), or in R (version 3.6.1, or higher, R Foundation for Statistical Computing, Vienna, Austria).

Data Availability

Software for running the task, as well as the data and analysis code for regenerating our main results and figures are available as a Zenodo release (https://doi.org/10.5281/zenodo.5513429) of a GitHub repository

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(https://github.com/erlichlab/gpstress).

Author Contributions

All authors contributed to the study design, interpretation of the results and writing the manuscript. Data collection and analyses were performed by E. Lukinova under the supervision of J. Erlich.

Acknowledgments

We acknowledge an outstanding NYU Shanghai undergraduate who helped us to start this project: Yuyue Wang. We acknowledge Brianna Fu from NYU, Danielle John from CUNY Hunter College and Shengjie Xu from NYU Shanghai who helped us to recruit Chinese general population participants and run sessions. We thank NYU Shanghai undergraduate students Xirui Zhao and Muzi Du who helped us to run ELISA on saliva samples. We acknowledge Prof. Wenshu Li who allowed and supported our use of NYU Shanghai Molecular Biology Lab for running ELISA. Funding was provided by National Science Foundation of China (NSFC-31750110461) International Young Scholar Grant and Shanghai Eastern Scholar Program (Shanghai Municipal Education Commission) Talents Young Scholar Grant to E.L.

References

American Psychological Association and others. Stress in America findings. *Retrieved November*, 30:2013, 2010.

Pragathi P. Balasubramani, V. Srinivasa Chakravarthy, Balaraman Ravindran, and Ahmed A. Moustafa. An extended reinforcement learning model of basal ganglia to understand the contributions of serotonin and dopamine in risk-based decision making, reward prediction, and punishment learning. *Frontiers in Computational Neuroscience*, 8, 2014. ISSN 1662-5188. doi: 10.3389/fncom.2014.00047. URL `https://www.frontiersin.org/articles/10.3389/fncom.2014.00047/full`.

Efrat Barel, Shosh Shahrabani, and Orna Tzischinsky. Sex Hormone/Cortisol Ratios Differentially Modulate Risk-Taking in Men and Women. *Evolutionary Psychology*, 15(1):147470491769733, January 2017. ISSN 1474-7049, 1474-7049. doi: 10.1177/1474704917697333. URL `http://journals.sagepub.com/doi/10.1177/1474704917697333`.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, 67(1), 2015. ISSN 1548-7660. doi: 10.18637/jss.v067.i01. URL `http://www.jstatsoft.org/v67/i01/`.

Paul-Christian Bürkner. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1):1–28, 2017. doi: 10.18637/jss.v080.i01.

Bob Carpenter, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 20:1–37, 2016.

Smarandita Ceccato, Brigitte M Kudielka, and Christiane Schwieren. Increased risk taking in relation to chronic stress in adults. *Frontiers in psychology*, 6, 2015.

Zheng Chen, Jifeng Li, Jing Zhang, Xue Xing, Wei Gao, Zuhong Lu, and Huihua Deng. Simultaneous determination of hair cortisol, cortisone and DHEAS with liquid chromatography–electrospray ionization-tandem mass spectrometry in negative mode. *Journal of Chromatography B*, 929:187–194, June 2013. ISSN 15700232. doi: 10.1016/j.jchromb.2013.04.026. URL `https://linkinghub.elsevier.com/retrieve/pii/S1570023213002365`.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, 24(4):385, December 1983. ISSN 00221465. doi: 10.2307/2136404. URL `http://www.jstor.org/stable/2136404?origin=crossref`.

Alain Cohn, Jan Engelmann, Ernst Fehr, and Michel André Maréchal. Evidence for countercyclical risk aversion: an experiment with financial professionals. *The American Economic Review*, 105(2):860–885, 2015.

Sandra Cornelisse, Vanessa Van Ast, Johannes Haushofer, Maayke Seinstra, and Marian Joels. Time-dependent effect of hydrocortisone administration on intertemporal choice. preprint, SSRN, 2013. URL `http://dx.doi.org/10.2139/ssrn.2294189`.

Carlos Cueva, R Edward Roberts, Tom Spencer, Nisha Rani, Michelle Tempest, Philippe N Tobler, Joe Herbert, and Aldo Rustichini. Cortisol and testosterone increase financial risk taking and may destabilize markets. *Scientific reports*, 5:11206, 2015.

Birk Diedenhofen and Jochen Musch. cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLOS ONE*, 10(4):e0121945, April 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0121945. URL `http://dx.plos.org/10.1371/journal.pone.0121945`.

Andrea Dietrich, Johan Ormel, Jan K Buitelaar, Frank C Verhulst, Pieter J Hoekstra, and Catharina A Hartman. Cortisol in the morning and dimensions of anxiety, depression, and aggression in children from a general population and clinic-referred cohort: an integrated analysis. The TRAILS study. *Psychoneuroendocrinology*, 38(8):1281–1298, 2013.

Carlo Faravelli and Stefano Pallanti. Recent life events and panic disorder. *Am J Psychiatry*, 146(5):622–6, 1989.

Scott H Frank and Stephen J Zyzanski. Stress in the clinical setting: the Brief Encounter Psychosocial Instrument. *The Journal of family practice*, 26(5):533–9, June 1988.

Michael Friendly. Corrgrams: Exploratory Displays for Correlation Matrices. *The American Statistician*, 56(4): 316–324, November 2002. ISSN 0003-1305, 1537-2731. doi: 10.1198/000313002533. URL `http://www.tandfonline.com/doi/abs/10.1198/000313002533`.

Elizabeth V. Goldfarb, Monja I. Froböse, Roshan Cools, and Elizabeth A. Phelps. Stress and Cognitive Flexibility: Cortisol Increases Are Associated with Enhanced Updating but Impaired Switching. *Journal of Cognitive Neuroscience*, 29(1):14–24, January 2017. ISSN 0898-929X, 1530-8898. doi: 10.1162/jocn_a_01029. URL `https://www.mitpressjournals.org/doi/abs/10.1162/jocn_a_01029`.

J González-Cabrera, M Fernández-Prada, C Iribar-Ibabe, and JM Peinado. Acute and chronic stress increase salivary cortisol: a study in the real-life setting of a national examination undertaken by medical graduates. *Stress*, 17(2):149–156, 2014.

Holly C Gooding, Carly E Milliren, S Bryn Austin, Margaret A Sheridan, and Katie A McLaughlin. Child abuse, resting blood pressure, and blood pressure reactivity to psychosocial stress. *Journal of pediatric psychology*, page jsv040, 2015.

Lisa A Goodman, Carole Corcoran, Kiban Turner, Nicole Yuan, and Bonnie L Green. Assessing traumatic event exposure: General issues and preliminary findings for the Stressful Life Events Screening Questionnaire. *Journal of traumatic stress*, 11(3):521–542, 1998.

R. Gow, S. Thomson, M. Rieder, S. Van Uum, and G. Koren. An assessment of cortisol analysis in hair and its clinical applications. *Forensic Science International*, 196(1-3):32–37, March 2010. ISSN 03790738. doi: 10.1016/j.forsciint.2009.12.040. URL `https://linkinghub.elsevier.com/retrieve/pii/S0379073809005416`.

Constance Hammen. Stress and depression. *Annu. Rev. Clin. Psychol.*, 1:293–319, 2005.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Allison G Harvey, Charlie Jones, and D Anne Schmidt. Sleep and posttraumatic stress disorder: a review. *Clinical Psychology Review*, 23(3):377–407, 2003.

Johannes Haushofer and Ernst Fehr. On the psychology of poverty. *Science*, 344(6186):862–867, May 2014. doi: 10.1126/science.1232491. URL `http://www.sciencemag.org/cgi/content/abstract/344/6186/862`. 00000.

Dirk H Hellhammer, Stefan Wüst, and Brigitte M Kudielka. Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, 34(2):163–171, 2009.

Thomas H Holmes and Richard H Rahe. The social readjustment rating scale. *Journal of psychosomatic research*, 11(2):213–218, 1967.

Lisa Huang, Amir Rattner, Han Liu, and Jeremy Nathans. How to draw the line in biomedical research. *eLife*, 2, March 2013. ISSN 2050-084X. doi: 10.7554/eLife.00638. URL `https://elifesciences.org/articles/00638`.

Warrick J Inder, Goce Dimeski, and Anthony Russell. Measurement of salivary cortisol in 2012–laboratory techniques and clinical indications. *Clinical endocrinology*, 77(5):645–651, 2012.

Narayanan Kandasamy, Ben Hardy, Lionel Page, Markus Schaffner, Johann Graggaber, Andrew S Powlson, Paul C Fletcher, Mark Gurnell, and John Coates. Cortisol shifts financial risk preferences. *Proc Natl Acad Sci U S A*, 111(9):3608–3613, March 2014. doi: 10.1073/pnas.1317908111. URL `http://www.pnas.org/cgi/doi/10.1073/pnas.1317908111`. 00000.

Jerker Karlén, Johnny Ludvigsson, Anneli Frostell, Elvar Theodorsson, and Tomas Faresjö. Cortisol in hair measured in young adults-a biomarker of major life stressors? *BMC Clinical Pathology*, 11(1):12, 2011.

KS Kendler, LM Karkowski, and CA Prescott. The assessment of dependence in the study of stressful life events: validation using a twin design. *Psychological medicine*, 29(06):1455–1460, 1999.

Lars Vedel Kessing, Esben Agerbo, and Preben B Mortensen. Does the impact of major stressful life events on the risk of developing depression change throughout life? *Psychological medicine*, 33(07):1177–1184, 2003.

Karolina M. Lempert, Anthony J. Porcelli, Mauricio R. Delgado, and Elizabeth Tricomi. Individual Differences in Delay Discounting Under Acute Stress: The Role of Trait Perceived Stress. *Frontiers in Psychology*, 3, July 2012. ISSN 1664-1078. doi: 10.3389/fpsyg.2012.00251. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3400439/`. 00037.

Karolina M. Lempert, Joseph T. McGuire, Danielle B. Hazeltine, Elizabeth A. Phelps, and Joseph W. Kable. The effects of acute stress on the calibration of persistence. *Neurobiology of Stress*, 8:1–9, February 2018a. ISSN 23522895. doi: 10.1016/j.ynstr.2017.11.001. URL `https://linkinghub.elsevier.com/retrieve/pii/S2352289517300310`.

Karolina M. Lempert, Joanna E. Steinglass, Anthony Pinto, Joseph W. Kable, and Helen Blair Simpson. Can delay discounting deliver on the promise of RDoC? *Psychological Medicine*, pages 1–10, August 2018b. ISSN 0033-2917, 1469-8978. doi: 10.1017/S0033291718001770. URL `https://www.cambridge.org/core/product/identifier/S0033291718001770/type/journal_article`.

Karolina Maria Lempert, David A. Wolk, and Joe Kable. Declarative memory, but not executive function, is associated with temporal discounting in older adults. preprint, PsyArXiv, September 2018c. URL `https://osf.io/vb743`.

Jennifer K. Lenow, Sara M. Constantino, Nathaniel D. Daw, and Elizabeth A. Phelps. Chronic and Acute Stress Promote Overexploitation in Serial Decision Making. *The Journal of Neuroscience*, 37(23):5681–5689, June 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3618-16.2017. URL `http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3618-16.2017`.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Jennifer S Lerner, Ye Li, and Elke U Weber. The financial costs of sadness. *Psychological science*, page 0956797612450302, 2012.

Arnold Stanley Linsky and Murray Arnold Straus. *Social stress in the United States: Links to regional patterns in crime and illness*. Auburn House, 1986.

R. Linz, T. Singer, and V. Engert. Interactions of momentary thought content and subjective stress predict cortisol fluctuations in a daily life experience sampling study. *Scientific Reports*, 8(1):15462, December 2018. ISSN 2045-2322. doi: 10.1038/s41598-018-33708-0. URL `http://www.nature.com/articles/s41598-018-33708-0`.

Silvia Lopez-Guzman, Anna B. Konova, and Paul W. Glimcher. Computational psychiatry of impulsivity and risk: how risk and time preferences interact in health and disease. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1766):20180135, February 2019. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2018.0135. URL `http://www.royalsocietypublishing.org/doi/10.1098/rstb.2018.0135`.

Evgeniya Lukinova and Jeffrey C Erlich. Quantifying the contribution of individual variation in timing to delay-discounting. *Scientific reports*, 11:18354, 2021. doi: 10.1038/s41598-021-97496-w.

Evgeniya Lukinova, Yuyue Wang, Steven F Lehrer, and Jeffrey C Erlich. Time preferences are reliable across time- horizons and verbal versus experiential tasks. *eLife*, page 27, February 2019. doi: https://doi.org/10.7554/eLife.39656.001.

Gediminas Luksys and Carmen Sandi. Neural mechanisms and computations underlying stress effects on learning and memory. *Current Opinion in Neurobiology*, 21(3):502–508, June 2011. ISSN 09594388. doi: 10.1016/j.conb.2011.03.003. URL `https://linkinghub.elsevier.com/retrieve/pii/S0959438811000432`.

Silvia U. Maier, Aidan B. Makwana, and Todd A. Hare. Acute Stress Impairs Self-Control in Goal-Directed Choice by Altering Multiple Functional Connections within the Brain's Decision Circuits. *Neuron*, 87(3): 621–631, August 2015. ISSN 08966273. doi: 10.1016/j.neuron.2015.07.005. URL `https://linkinghub.elsevier.com/retrieve/pii/S0896627315006273`.

Jennifer Meggs, J Golby, CJ Mallett, DF Gucciardi, and Remco CJ Polman. The Cortisol Awakening Response and Resilience in Elite Swimmers. *International journal of sports medicine*, 37(02):169–174, 2016.

Gregory E Miller, Edith Chen, and Eric S Zhou. If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. *Psychological bulletin*, 133(1):25, 2007.

Sendhil Mullainathan and Eldar Shafir. *Scarcity: Why having too little means so much*. Macmillan, 2013.

Jonathan W Peirce. PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, 162(1): 8–13, 2007.

Vicente Prado-Gascó, Usue de la Barrera, Sandra Sancho-Castillo, José Enrique de la Rubia-Ortí, and Inmaculada Montoya-Castilla. Perceived stress and reference ranges of hair cortisol in healthy adolescents. *PLOS ONE*, 14(4):e0214856, April 2019. ISSN 1932-6203. doi: 10.1371/journal.pone.0214856. URL `https://dx.plos.org/10.1371/journal.pone.0214856`.

Marita Pruessner, Dirk H Hellhammer, Jens C Pruessner, and Sonia J Lupien. Self-reported depressive symptoms and stress levels in healthy young men: associations with the cortisol response to awakening. *Psychosomatic Medicine*, 65(1):92–99, 2003.

Robbie Schepers and C Rob Markus. Gene$\times$ cognition interaction on stress-induced eating: Effect of rumination. *Psychoneuroendocrinology*, 54:41–53, 2015.

Anuj K Shah, Sendhil Mullainathan, and Eldar Shafir. Some consequences of having too little. *Science*, 338 (6107):682–685, 2012.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.18.460891; this version posted September 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Peter Sokol-Hessner, Candace M. Raio, Sarah P. Gottesman, Sandra F. Lackovic, and Elizabeth A. Phelps. Acute stress does not affect risky monetary decision-making. *Neurobiology of Stress*, 5:19–25, December 2016. ISSN 2352-2895. doi: 10.1016/j.ynstr.2016.10.003. 00005.

Tobias Stalder, Susann Steudte-Schmiedgen, Nina Alexander, Tim Klucken, Annika Vater, Susann Wichmann, Clemens Kirschbaum, and Robert Miller. Stress-related and basic determinants of hair cortisol in humans: A meta-analysis. *Psychoneuroendocrinology*, 77:261–274, March 2017. ISSN 03064530. doi: 10.1016/j.psyneuen.2016.12.017. URL `https://linkinghub.elsevier.com/retrieve/pii/S0306453016306114`.

Taiki Takahashi. Cortisol levels and time-discounting of monetary gain in humans. *Neuroreport*, 15(13): 2145–2147, September 2004. ISSN 0959-4965. 00063.

Agnieszka Tymula, Lior A Rosenberg Belmaker, Lital Ruderman, Paul W Glimcher, and Ifat Levy. Like cognitive function, decision making across the life span shows profound age-related changes. *Proceedings of the National Academy of Sciences*, 110(42):17143–17148, 2013.

Berry J. van Holland, Monique H. W. Frings-Dresen, and Judith K. Sluiter. Measuring short-term and long-term physiological stress effects by cortisol reactivity in saliva and hair. *International Archives of Occupational and Environmental Health*, 85(8):849–852, November 2012. ISSN 0340-0131, 1432-1246. doi: 10.1007/s00420-011-0727-3. URL `http://link.springer.com/10.1007/s00420-011-0727-3`.

Barbara Vanaelst, Inge Huybrechts, Karin Bammann, Nathalie Michels, Tineke Vriendt, Krishna Vyncke, Isabelle Sioen, Licia Iacoviello, Kathrin Günther, Denes Molnar, Lauren Lissner, Noellie Rivet, Jean-Sebastien Raul, and Stefaan Henauw. Intercorrelations between serum, salivary, and hair cortisol and child-reported estimates of stress in elementary school girls: Cortisol intercorrelations and childhood stress. *Psychophysiology*, pages n/a–n/a, July 2012. ISSN 00485772. doi: 10.1111/j.1469-8986.2012.01396.x. URL `http://doi.wiley.com/10.1111/j.1469-8986.2012.01396.x`.

Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432, September 2017. ISSN 0960-3174, 1573-1375. doi: 10.1007/s11222-016-9696-4. URL `https://link.springer.com/article/10.1007/s11222-016-9696-4`. 00084.

Lisa J. Weckesser, Friedericke Dietz, Kornelius Schmidt, Juliane Grass, Clemens Kirschbaum, and Robert Miller. The psychometric properties and temporal dynamics of subjective stress, retrospectively assessed by different informants and questionnaires, and hair cortisol concentrations. *Scientific Reports*, 9(1):1098, December 2019. ISSN 2045-2322. doi: 10.1038/s41598-018-37526-2. URL `http://www.nature.com/articles/s41598-018-37526-2`.

Quan Zhang, Zheng Chen, Shenghuo Chen, Tian Yu, Juxia Wang, Weiwen Wang, and Huihua Deng. Correlations of hair level with salivary level in cortisol and cortisone. *Life Sciences*, 193:57–63, January 2018. ISSN 00243205. doi: 10.1016/j.lfs.2017.11.037. URL `https://linkinghub.elsevier.com/retrieve/pii/S0024320517306148`.