The effects of acute and chronic stress on choice consistency

Felix J. Nitsch, Manuela Sellitto, and Tobias Kalenscher

Comparative Psychology, Heinrich-Heine-University Düsseldorf, Germany

Wordcount: 4705

Author Note

Felix J. Nitsch https://orcid.org/0000-0002-7832-7498

Manuela Sellitto https://orcid.org/0000-0002-3987-7526

Tobias Kalenscher https://orcid.org/0000-0002-0358-9020

Correspondence concerning this article should be addressed to Felix J. Nitsch, Comparative Psychology, Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany. Email: felix.nitsch@hhu.de

The research was supported by a grant from the Deutsche Forschungsgemeinschaft (Bonn, Germany) to Tobias Kalenscher (grant-ID: DFG-KA: DFG-KA 2675/4-3).

We thank Adam Schweda for helpful comments on the study design as well as assistance with the data collection. We thank Johanna Falk, Anne Schmittner, Jochen Hoog, Alina Keßler and Paul Kramer for assistance with data collection.

Abstract

Important decisions are often made under some degree of stress. It is now well-established that acute stress affects preferences and priorities in our decisions. However, it is hard to make a general case on the net impact of stress on decision-making quality in a normative sense as evidence for or against a direct effect of stress on decision-making quality is sparse. Here we use the revealed preference framework of choice consistency to investigate decision-making quality without the assumption of an objectively correct choice. Specifically, we tested whether acute stress influences choice consistency in a time dependent fashion. A sample of 144 participants solved a food choice task before, immediately after and in the aftermath of the Trier Social Stress Test (TSST) or a matched control procedure. We confirmed the effectiveness of our stress manipulation via an array of subjective and physiological stress measures. Using Bayesian statistics, we found strong evidence against an effect of acute stress on choice consistency. However, we found exploratory evidence for a negative association of self-reported chronic stress and choice consistency. We discuss our results in the context of previous findings of stress effects on choice consistency and preference changes.

Keywords: Stress, choice consistency, rationality, cortisol, revealed preference

The effects of acute and chronic stress on choice consistency

Important decisions are often made under some degree of stress. In an increasingly fast-paced economy, managers have to decide whether to keep or kill a project with high financial and personal impact. Doctors have to select the most promising therapy when lives are at stake. Air traffic controllers have to declare a flight safe to start, evaluating a complex multi-factorial system. It is crucial to understand how stress might affect the ability to make good decisions.

In modern economics, decision-making quality is not defined by the specific direction of preferences but instead by bounds on the preference structure in the form of consistency principles (Sugden, 1991). Revealed preference theory (Houthakker, 1950; Samuelson, 1938; Varian, 1982) requires decision-makers to have a well-defined and stable set of preferences, as well as to act cost-efficiently and consistently upon their fulfillment, which is mathematically equivalent with optimizing decision-making (Afriat, 1973). Importantly, revealed preference optimality does not require objectively, or at least normatively correct preferences (Choi et al., 2014). The notion of decision-making quality as choice consistency can be illustrated by an example:

Assume that a manager chooses to endorse a project A instead of another project B when both require the same budget. The same manager must, everything else equal, also endorse project A when it is cheaper than project B. That is, because from the former decision we learn that the manager either prefers project A or is indifferent and chose randomly (because both projects were equally expensive). Now the necessity of choosing project A in the latter decision can be proven by exhaustion: If the manager prefers project A over B, they should always choose project A as long as it is at least as cheap as project B. If the manager is indifferent between both projects, they should always choose the cheaper one, which is project A in the second decision, in order to act cost-efficiently. Also, transitivity must hold for these *revealed preferences*. For example, if the manager chose project A over B and project B over C, always assuming equal project costs, then they should not choose project C over project A, when project A is cheaper. In our example, it is irrelevant

whether and why the manager prefers project A or B or even is indifferent, as long as they are, everything else equal, consistent in their decisions.

Among several other factors, stress has been shown to affect the preferences and priorities in our decisions concerning food (Habhab et al., 2009; Maier et al., 2015; Zellner et al., 2006), other people (Margittai et al., 2015; Margittai, Van Wingerden, et al., 2018; Schweda et al., 2019; Vinkers et al., 2013; Von Dawans et al., 2012), risky prospects (Margittai, Nave, et al., 2018; Starcke & Brand, 2016), as well as the timing of financial returns (Cornelisse et al., 2013; Kimura et al., 2013; Riis-Vestergaard et al., 2018). These behavioral effects are paralleled by resource reallocation across large-scale brain networks during the stress response (Hermans et al., 2011, 2014). Immediately after the stressful event, catecholamines upregulate vigilance-related functions while prefrontal functions including executive control are inhibited (Arnsten, 2009; Arnsten et al., 2012; Hermans et al., 2011; Qin et al., 2009). With slight temporal offset, this pattern is enhanced by rapid, non-genomic corticosteroid effects (Hermans et al., 2011; Tasker et al., 2006). In the aftermath of stress genomic, corticosteroid effects are working to reestablish homeostasis (De Kloet et al., 2005), but repeated stress exposure slows the recovery of homeostasis after stressful events (Cameron & Schoenfeld, 2019).

Hence, stress affects preferences in various choice domains. However, as argued above, such preference shifts are not sufficient to show decreased decision quality. For example, the increased preference for calorie-dense food and immediate rewards under stress might simply reflect the attempt to replenish depleted energy resources; hence, choices for high-caloric food might very well be consistent with the decision-maker's preferences.

A more convincing argument for the possibility that stress could impair decision-making quality stems from the observation that underlying cognitive decision-making abilities deteriorate under stress. For example, acute stress, or stress hormone action, has been shown to impair working-memory, problem-solving abilities, self-control, and cognitive reflection (Beilock & DeCaro, 2007; Maier et al., 2015; Margittai et al., 2016; Qin et al., 2009). All of these psychological functions

have been linked to decision-making (Brand et al., 2005, 2014; Frederick, 2005; Gathmann et al., 2014). In addition, repeated exposure to stress (i.e. chronic stress) has been robustly linked to decreased mental health, particularly depression, an increased number of cognitive failures as well as neurotoxic effects in general (Cameron & Schoenfeld, 2019; Linden et al., 2005; Marin et al., 2011). Animal research has provided evidence that chronic stress affects the ability to perform actions based on their consequences (Dias-Ferreira et al., 2009), which is an integral feature of utilitarian decision-making. Further, chronic stress in animals has been linked to increased high-cost/high-reward choices over a wide range of options (Friedman et al., 2017). Still, it does not follow immediately, that decision-making quality must be impaired if certain cognitive abilities are. Through means of resource reallocation, deficiencies in one ability might be compensated by an upsurge in another.

Overall, evidence for or against a direct effect of stress on decision-making quality is sparse, with only one study testing the influence of an acute physical stressor on choice consistency in decisions among risky prospects (Cettolin et al., 2019). Here, we asked whether acute stress impairs decision-making quality in the revealed preference sense defined above. That is, we asked whether acute stress increases the tendency to make decisions against one's previously revealed preferences. Previous research suggests that the immediate effects of the stress response impairs functions important for decision-making such as executive control, while the aftermath of the stress response enhances them. Therefore, we tested the following hypothesis:

Hypothesis: Acute stress influences choice consistency in a time dependent fashion similar to the effects on executive functions. Specifically, we expect choice consistency to be impaired immediately after stressor offset, but not in the aftermath of a stressor.

Further, cognitive functioning is impaired under chronic stress exposure. While the investigation of chronic stress was not the main objective of the current study, we assessed self-reported chronic stress exposure to explore its influence on choice consistency

Explorative question: Is chronic stress exposure associated with impaired choice consistency?

Methods

Sample characteristics and inclusion criteria

We included 144 participants (76 females). Participants did not have formal psychological or economic education, were 18 – 40 years old, non-smokers and did not take medication that could have influenced their corticosteroid levels. Women were not taking oral contraceptives. Similar to previous studies (Margittai et al., 2015; Schweda et al., 2019), participants had to refrain from drinking alcohol and sexual activities 24 hours, caffeine four hours and eating/drinking (except water) two hours before the experiment. Table 1 provides summary statistics of the sample characteristics.

Procedure

Before the experimental session, all participants were screened for the inclusion criteria, rated their liking of several fruits/vegetables and sweet/salty snacks on a 5-point Likert scale online, and responded to various psychological trait measures via an online questionnaire (see trait measures).

All experimental sessions took place from 3 p.m. to 6 p.m. to control for circadian variations of hormonal levels (Van Cauter, 1990). Participants were assigned to the two experimental conditions pseudo-randomly. In the lab, participants completed the behavioral task at three different time points (figure 1): on arrival (*baseline time point*), within ten minutes after the stress/control procedure offset (*early time point*), and 90 minutes after stress/control procedure offset (*late time point*). Between the early and the late time point, participants were allowed to read magazines but not to talk or to use their phones.

Stress manipulation

In order to experimentally induce acute psychosocial stress, we used the group version of the Trier Social Stress Test (TSST-G): During the 20 min long TSST-G procedure, groups of four women or men were asked to carry out a fictional job interview (3 minutes per participant) and a mental arithmetic task (deducting from a four-digit number in steps of 16; 1.5 minutes per participant) in front of an evaluative panel of experts while being videotaped. The panel consisted of a female and a male panelist. For women, the male panelist took up the active role and the female panelist, vice-versa, for men. Panelist wore white lab coats. They closely observed the participants, interrupted them and took notes of their performance, and did not give any verbal or non-verbal positive feedback. We used a matched control procedure for the control group (von Dawans et al., 2011), in which groups of three to four women or men were instructed to tell a story about a good friend and complete the mental arithmetic task simultaneously. To compensate for the shorter, simultaneous procedure, participants were instructed to read magazines in between tasks. They were neither videotaped nor directly observed by the panel, removing the component of social evaluation but keeping other aspects of the procedure as similar as possible. The panel also did not were lab coats in the control condition. The TSST is considered the gold-standard for stress-induction in humans and has been validated extensively (Het et al., 2009; Kirschbaum et al., 1993; Kudielka & Kirschbaum, 2005; Rohleder et al., 2004) and successfully established in our lab (Margittai et al., 2015; Schweda et al., 2019).

Physiological and subjective stress measures

To assess the effectiveness of the stress manipulation, we collected a saliva sample before and after each behavioral task time point as well as during the stress/control procedure and measured cortisol and alpha-amylase, a measure of noradrenergic activity (Nater & Rohleder, 2009), for a total of 7 measurements across the experiment (figure 1, upper panel). Saliva samples were collected using Salivette devices (Sarstedt, Germany) containing a cotton wool swab that participants had to lightly chew on for 60 seconds to allow the swab to fill with saliva. Samples were

frozen and stored at −20 degree Celsius and subsequently sent to the Dresden LabService GmbH for cortisol and alpha-amylase measurement. Both cortisol and alpha-amylase concentrations were determined using a luminescent immunoassay (IBL International, Hamburg). For 16 samples analysis of cortisol and/or alpha-amylase was not possible due to insufficient saliva or contamination with blood. Furthermore, we assessed the heartrate of the participants at baseline and during the stress/control procedure with commercial wristbands (Polar A370) and their current affect with the Positive and Negative Affect Scale (Krohne et al., 1996) and four visual analogue scales (items: *stressed*, *ashamed*, *insecure*, *self-secure;* 100mm scale). We assessed the natural chronic stress exposure of our participants prior to the experiment with the Trier Inventory of Chronic Stress (TICS; Petrowski et al., 2012). The questionnaire comprises 57 items, which ask for 9 types of stress. These include work overload, social overload, pressure to succeed, dissatisfaction with the job, excessive demands at the job, lack of social recognition, social tension, social isolation and chronic worries. Each item must be scored on a five-level scale from 0 to 4. These values are added together to form the total score. A higher score corresponds to higher chronic stress exposure.

Trait measures

Besides the TICS, we administered the following questionnaires prior to the experiment to exclude the possibility of spurious trait differences between stress and control participants: the 10-item version of the Big-Five-Inventory (BFI-10; Rammstedt & John, 2007), the behavioral inhibition & activation scale (BIS/BAS; Strobel et al., 2001), the quick delay questionnaire (QDQ; Clare et al., 2010), the social desirability scale (SDS; Stöber, 1999) and the morningness-eveningness-scale (MEQ; Griefahn et al., 2001). Furthermore, we assessed verbal intelligence via the multiple-choice vocabulary test (MWT-B; Lehrl et al., 1995) at the end of the laboratory session, as participants could have easily cheated in an online version.

Food choice task

We deployed a standard food choice task similar to the one used by Harbaugh et al. (2001) and Chung et al. (2017). We specifically decided for a food choice task over other choice domains

(i.e. risk or time preferences) to minimize task complexity and, thus, ensure that any inconsistencies would not arise simply from misunderstanding of the task. This was especially important as we recruited a sample that was naïve to economic theory.

In each trial, participants had to choose one out of a set of two to seven snack bundles. Each snack bundle consisted of specific amounts of a sweet or salty snack and a fruit or vegetable (see figure 1, lower panel). The snacks for both categories were selected to be similarly attractive according to the previously provided online ratings of the participants. At each of the three time points, participants had to make choices in 11 trials. The 11 trials for each time point were randomly sampled from a collection of 22 possible trials (see supplemental figure S1). The sampling procedure was implemented to reduce interdependency of the answers of each participant for subsequent time points, while keeping the presented bundle size out of satiation range. At the end of the experiment one random decision of each participant was implemented and the participant received the corresponding snack bundle. Using simulated data of 10.000 uniform-randomly deciding virtual decision-makers, we ensured that our design was sufficiently powerful to detect inconsistent behavior (see below for a quantitative definition of inconsistency): $Bronar's\ Power = 0.972$ (Bronars, 1987). Supplemental figure S2 summarizes the results of our simulation study.

For 8 participants, no or incomplete food choice data were saved due to a technical failure of the experimental hardware.

Analysis pipeline

For all analyses, we used a Bayesian framework of inference. Bayesian statistics allows us to express confidence in that a parameter is within a certain range, to extend parameter estimation naturally for complicated models, and to express evidence for or against hypotheses on a continuous scale (Wagenmakers et al., 2018). The latter point is especially relevant if we want to make statements about the absence of a possible effect, not only the absence of evidence.

Decision-making quality analysis

Instead of a normative evaluation of preferences, we defined decision quality as the consistency with which participants pursued their individual preferences regardless of their direction. That is, we assessed choice consistency, not choice content. Notably, this implies that choices of high-caloric, high-fat and high-sugar foods could potentially be considered rational, as long as such choices were consistent.

To this means, we calculated the Critical Cost Efficiency Index (Afriat, 1972; Varian, 1991) for each participant per time point. The Critical Cost Efficiency Index (CCEI) is the most widely used measure of rationality or decision-making quality in the framework of revealed preference theory. The CCEI exploits the fact that, in theory, inconsistent behavior is not cost-efficient, which can be illustrated by an example:

Assume, that from a previous choice we learn that 5 chocolate bars are at least as valuable to a decision-maker as 5 pieces of orange. Let's further assume that at another point in time, the decision-maker still buys 5 orange pieces although they cost 10% more than 5 chocolate bars. As we know that 5 chocolate bars are at least as good as 5 orange pieces, they spent at least 10% too much money for the obtain valued and, conversely, showed a 90% cost efficiency.

This would be denoted by a CCEI of 0.9. A CCEI of 1 denotes 100% cost efficiency and perfect consistency. The index approaches zero as the behavior becomes more inconsistent (for a different interpretation of the CCEI see also Nitsch & Kalenscher, 2020).

Hormonal analysis

Values were not subjected to any data transformations (Feng et al., 2014). We excluded one data point from our cortisol data (control group, measurement "TSST-G/control", VPN111) which deviated more than 20 standard deviations from the grand mean of our cohort and was more than 60 times higher than the median reference value for females her age, controlling for wake-up time (Miller et al., 2016).

Statistical model

We used a mixed-factorial design with time point (baseline, early, late) as a within-subject factor and experimental group (stress, control) as a between-subject factor. In statistical terms we used a mixed-factor ANOVA-style model including time point as within-subject factor and experimental group as between-subject factor, and the CCEI as dependent variable. Our main interest lies in the possible interaction of both factors as a statistical representation of the time-dependent influence of acute stress on choice consistency (see hypotheses). Hence, our interpretation of the data will be guided by two model comparisons:

1. The evidence for the full model including both main effects as well as the interaction effect compared to a null model (including only a subject-level intercept).
2. The evidence for the full model including both main effects as well as the interaction effect compared to a reduced model only containing both main effects but not the interaction effect.

As previous evidence on the time-dependent effect of acute stress on choice consistency is limited, we will use uninformative default priors for our model parameters (Rouder et al., 2012). We used structurally similar models for our manipulation checks (physiological and subjective stress measures). To test for an association of self-reported chronic stress and choice consistency we used a Bayes Factor test for correlations. Lastly, to test for absence of trait differences between stress and control participants we used pair-wise Bayesian t-tests contingency tables for the assessed questionnaire scores and demographic variables.

Data Collection Plan

Our data collection plan was based on a Bayesian stopping rule: We planned to collect data until we had reached a Bayes factor of $BF \geq 10 \vee BF \leq 0.1$ for all manipulation checks as well as our hypothesis, but at least for one year, to ensure interpretability of our analysis. Evidence for our explorative question was secondary to our sample size rationale and explorative results should be interpreted accordingly.

Results

Demographic variables and trait measures

We did not find strong evidence for any difference of the two experimental groups regarding all demographic variables and trait measures considered (see table 1). Accordingly, we can consider our randomization of participants into experimental conditions successful.

Cortisol and alpha-amylase

The data for both hormonal markers indicated that our stress manipulation was successful. The full model including both main effects and the interaction effect was much more likely than the null model given the data for both cortisol ($BF_{10} > 100$) and amylase ($BF_{10} > 100$), which is conventionally interpreted as extreme evidence (Jeffreys, 1998). Similarly, we found extreme evidence for the inclusion of the interaction effect by comparing the full model to reduced model for both cortisol ($BF_{10} > 100$) and amylase ($BF_{10} > 100$). Figure 2 shows a plot of the data. In summary, we found extreme evidence for the effectiveness of our stress manipulation in increasing the levels of salivary cortisol and alpha-amylase.

Heartrate

The effectiveness of our stress manipulation was corroborated by the heartrate data. The full model including both main effects and the interaction effect was much more likely than the null model given the data ($BF_{10} > 100$), which is conventionally interpreted as extreme evidence (Jeffreys, 1998). Similarly, we found extreme evidence for the inclusion of the interaction effect by comparing the full model to reduced model ($BF_{incl} > 100$). Figure 3 shows a plot of the data. In summary, we found extreme evidence for the effectiveness of our stress manipulation as operationalized by the heartrate.

Self-reported affect

Lastly, stress dynamically influenced positive and negative affect, as well as self-reported stress, shame and insecurity (but not security): We found strong or extreme evidence for the

effectiveness of our stress manipulation as operationalized by 5 out of 6 of the self-report scales (see supplemental table). Figure 4 shows a plot of the data.

Hypothesis: Acute stress and choice consistency

Our behavioral results indicated that participants acutely stressed by the TSST-G did not dynamically differ in their choice consistency compared to not-stressed participants. The null model was 24 times more likely than the full model given the data ($BF_{10} = 24.23 \pm 2.27\%$), which conventionally can be interpreted as strong evidence for the absence of an effect (Jeffreys, 1998). Similarly, we found strong evidence against the inclusion of the interaction effect by comparing the full model to a reduced model ($BF_{incl} = 15.26 \pm 4.74\%$; see figure 5, upper panel). To corroborate the robustness of our results, we repeated the analysis including self-reported sex and age as covariates into all of our models. Results remained qualitatively unchanged (BF_{10} =23.61 ±3.77%; $BF_{incl} = 14.92 \pm 4.19\%$). In summary, we found strong evidence against our hypothesis. That is, participants in the TSST-G (acute stress) condition did not dynamically differ in their choice consistency compared to not-stressed participants but, instead, showed comparable consistency levels across all time points of the experiment.

Exploration: Chronic stress and choice consistency

We found exploratory evidence that choice consistency decreased with higher chronic stress exposure (see figure 5, lower panel). A negative relationship of chronic stress and choice consistency was 9 times more likely than a null relationship given the data ($BF = 9.03$), which conventionally can be interpreted as moderate evidence (Jeffreys, 1998). The posterior estimate of the association under the assumption that it is non-zero was $\rho = -0.21, CI_{95\%} = [-0.36, -0.05]$. This, tentatively, suggests that chronically stressed participants showed reduced choice consistency, and the level of consistency decreased with increasing levels of chronic stress.

Discussion

Evidence for or against a direct effect of stress on decision-making quality is sparse, with only one other study testing the influence of a physical stressor on choice consistency in decisions among risky prospects (Cettolin et al., 2019). Here, we conducted an experimental test on the influence of a well-established social stress protocol on choice consistency in a food choice task. We specifically decided for a food choice task over other choice domains (i.e. risk or time preferences) to minimize task complexity and, thus, ensure that any inconsistencies would not arise simply from misunderstanding of the task. We corroborated the effectiveness of our stress induction using multiple subjective and physiological stress measures. Results showed strong evidence against a temporally dynamic effect of acute stress on choice consistency: both experimental groups showed comparable consistency levels over all time points of the experiment. We conclude that decision quality does not deteriorate under acute stress. Further, we explored the relationship of self-reported chronic stress and choice consistency. Interestingly, our results indicated that higher levels of self-reported chronic stress were associated with lower choice consistency. Hence, decision quality might be impaired with increasing levels of chronic stress. This explorational result should be tested more rigorously in a future, confirmatory study.

Our results are in line with the findings of Cettolin et al. (2019), who did not find a significant effect of an acute physical stressor on choice consistency both shortly after stressor offset and in the aftermath of the stressor. However, although Cettolin et al. (2019) used a sufficiently large sample size and a statistically well-powered design, they did not explicitly quantify evidence for the null hypothesis. Using Bayesian statistics, we were able to provide a statistically well-founded confirmation of Cettolin and colleagues' conclusions. Further, we used a variant of the TSST as our stress protocol, which is generally considered the gold-standard for stress-induction in humans and has been validated extensively (Het et al., 2009; Kirschbaum et al., 1993; Kudielka & Kirschbaum, 2005; Rohleder et al., 2004). Our results showed that the null findings of Cettolin et al. (2019) generalize to a stress protocol with a social evaluation component, which is often regarded essential

for the human stress response. Extending the findings of Cettolin and colleagues, we find explorative evidence that we find that chronic stress on the other hand does affect choice consistency.

At first glance, these results seem to contradict previous results of dynamic preference shifts in response to stress, especially in the domain of dietary decisions (Habhab et al., 2009; Maier et al., 2015; Zellner et al., 2006). As stated above, choice consistency in the sense of revealed preference theory requires decision-makers to have a well-defined and stable set of preferences (Afriat, 1973). However, with our experimental procedure we ensured that each run of the food choice task for itself was completed within what we assume to be distinct time windows of the stress response (see figure 1). Thus, while we cannot exclude that preferences change between these time windows, we found strong evidence that choice consistency of stressed participants is not affected within these time windows as compared to not stressed participants. This is in so far intuitive, as altered preferences in response to a stressful event may very well be adaptive depending on the decision context, while a decreased level of cost-efficiency and choice consistency is rarely advantageous. Future research should challenge our results by investigating decision-making across hormonal time windows, e.g. by explicitly incorporating temporal dynamics in the modelling approach via leave-one-out measures of choice consistency (Kurtz-David et al., 2019).

It is important to point out that in the current study, our main focus was the effect of an acute stressful event on choice consistency in an incentivized choice task. While it is a potentially reassuring result that choice consistency is not impaired immediately following acute stress, or in its aftermath, we find tentative evidence that chronic stress does affect choice consistency. In line with this, animal research has found evidence that chronic stress affects the ability to perform actions based on their consequences (Dias-Ferreira et al., 2009), which is an integral feature of utilitarian decision-making. Another recent animal study found that chronic stress led to increased high-cost/high-reward choices over a wide range of options (Friedman et al., 2017). While the study could not determine whether this effect was driven by a reduced sensitivity to consequences, the authors argue that the cost-benefit integration was impaired in chronically stressed animals. Most likely due

to methodological constraints in the investigation of chronic stress, evidence in human subjects is sparser. One study found that chronic stress in jockeys was related to impaired decisions in relatively simple attention and reaction time tasks (Landolt et al., 2017). The authors note that impairments, descriptively, increased with task complexity. More research is necessary to confirm whether chronic stress, in contrast to acute stress, impairs choice consistency, possibly via a reduced sensitivity to action outcomes. To this means, for example, the contemporary COVID-19 pandemic provides a unique opportunity to investigate the influence of increased chronic stress exposure on choice consistency in the general population.

A last concern that we want to address here is generalizability regarding the presentational structure of the choice problem and, related to that, the conceptualization of choice consistency. The investigation of choice consistency has increasingly diversified over the last decades and there exist several, prima facie equally valid operationalizations of choice consistency or, in a wider sense, rationality (Rieskamp et al., 2006). The here utilized revealed preference framework of choice consistency is a necessary and sufficient condition for as-if utility maximization (Afriat, 1973) and, thus, does have theoretical appeal. However, future research of choice consistency should address whether violations of choice consistency are specific to certain axioms, e.g. stochastic dominance, transitivity or regularity. Previous, seemingly contradictory findings in the field of choice consistency might be reconciled when using a stringent theoretical framework of choice consistency.

References

Afriat, S. N. (1972). Efficiency Estimation of Production Functions. *International Economic Review*, *13*(3), 568–598. JSTOR. https://doi.org/10.2307/2525845

Afriat, S. N. (1973). On a system of inequalities in demand analysis: An extension of the classical method. *International Economic Review*, 460–472. https://doi.org/10.2307/2525934

Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, *10*(6), 410.

Arnsten, A. F., Wang, M. J., & Paspalas, C. D. (2012). Neuromodulation of thought: Flexibilities and vulnerabilities in prefrontal cortical network synapses. *Neuron*, *76*(1), 223–239.

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(6), 983–998. https://doi.org/10.1037/0278-7393.33.6.983

Brand, M., Fujiwara, E., Borsutzky, S., Kalbe, E., Kessler, J., & Markowitsch, H. J. (2005). Decision-Making Deficits of Korsakoff Patients in a New Gambling Task With Explicit Rules: Associations With Executive Functions. *Neuropsychology*, *19*(3), 267–277. https://doi.org/10.1037/0894-4105.19.3.267

Brand, M., Schiebener, J., Pertl, M.-T., & Delazer, M. (2014). Know the risk, take the win: How executive functions and probability processing influence advantageous decision making under risk conditions. *Journal of Clinical and Experimental Neuropsychology*, *36*(9), 914–929. Scopus. https://doi.org/10.1080/13803395.2014.955783

Bronars, S. G. (1987). The power of nonparametric tests of preference maximization. *Econometrica: Journal of the Econometric Society*, 693–698. https://doi.org/10.2307/1913608

Cameron, H. A., & Schoenfeld, T. J. (2019). *Behavioral and Structural Adaptations to Stress*. 21.

Cettolin, E., Dalton, P. S., Kop, W. J., & Zhang, W. (2019). Cortisol meets GARP: The effect of stress on economic rationality. *Experimental Economics*. https://doi.org/10.1007/s10683-019-09624-z

Choi, S., Kariv, S., Müller, W., & Silverman, D. (2014). Who Is (More) Rational? *American Economic Review*, *104*(6), 1518–1550. https://doi.org/10.1257/aer.104.6.1518

Chung, H.-K., Tymula, A., & Glimcher, P. (2017). The Reduction of Ventrolateral Prefrontal Cortex Gray Matter Volume Correlates with Loss of Economic Rationality in Aging. *The Journal of Neuroscience*, *37*(49), 1171–17. https://doi.org/10.1523/JNEUROSCI.1171-17.2017

Clare, S., Helps, S., & Sonuga-Barke, E. J. (2010). The quick delay questionnaire: A measure of delay aversion and discounting in adults. *ADHD Attention Deficit and Hyperactivity Disorders*, *2*(1), 43–48.

Cornelisse, S., Van Ast, V., Haushofer, J., Seinstra, M., & Joels, M. (2013). *Time-dependent effect of hydrocortisone administration on intertemporal choice*.

De Kloet, E. R., Joëls, M., & Holsboer, F. (2005). Stress and the brain: From adaptation to disease. *Nature Reviews Neuroscience*, *6*(6), 463.

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., Costa, R. M., & Sousa, N. (2009). Chronic Stress Causes Frontostriatal Reorganization and Affects Decision-Making. *Science*, *325*(5940), 621. https://doi.org/10.1126/science.1171203

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry*, *26*(2), 5.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Friedman, A., Homma, D., Bloem, B., Gibb, L. G., Amemori, K., Hu, D., Delcasso, S., Truong, T. F., Yang, J., Hood, A. S., Mikofalvy, K. A., Beck, D. W., Nguyen, N., Nelson, E. D., Toro Arana, S. E., Vorder Bruegge, R. H., Goosens, K. A., & Graybiel, A. M. (2017). Chronic Stress Alters Striosome-

Circuit Dynamics, Leading to Aberrant Decision-Making. *Cell*, *171*(5), 1191-1205.e28. https://doi.org/10.1016/j.cell.2017.10.017

Gathmann, B., Pawlikowski, M., Schöler, T., & Brand, M. (2014). Performing a secondary executive task with affective stimuli interferes with decision making under risk conditions. *Cognitive Processing*, *15*(2), 113–126. Scopus. https://doi.org/10.1007/s10339-013-0584-y

Griefahn, B., Künemund, C., Bröde, P., & Mehnert, P. (2001). Zur Validität der deutschen Übersetzung des Morningness-Eveningness-Questionnaires von Horne und Östberg: The Validity of a German Version of the Morningness-Eveningness-Questionnaire Developed by Horne and Östberg. *Somnologie*, *5*(2), 71–80.

Habhab, S., Sheldon, J. P., & Loeb, R. C. (2009). The relationship between stress, dietary restraint, and food preferences in women. *Appetite*, *52*(2), 437–444.

Harbaugh, W. T., Krause, K., & Berry, T. R. (2001). GARP for kids: On the development of rational choice behavior. *American Economic Review*, *91*(5), 1539–1545. https://doi.org/10.1257/aer.91.5.1539

Hermans, E. J., Henckens, M. J., Joëls, M., & Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in Neurosciences*, *37*(6), 304–314.

Hermans, E. J., van Marle, H. J., Ossewaarde, L., Henckens, M. J., Qin, S., van Kesteren, M. T., Schoots, V. C., Cousijn, H., Rijpkema, M., & Oostenveld, R. (2011). Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. *Science*, *334*(6059), 1151–1153.

Het, S., Rohleder, N., Schoofs, D., Kirschbaum, C., & Wolf, O. T. (2009). Neuroendocrine and psychometric evaluation of a placebo version of the 'Trier Social Stress Test'. *Psychoneuroendocrinology*, *34*(7), 1075–1086. https://doi.org/10.1016/j.psyneuen.2009.02.008

Houthakker, H. S. (1950). Revealed preference and the utility function. *Economica*, *17*(66), 159–174. https://doi.org/10.2307/2549382

Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.

Kimura, K., Izawa, S., Sugaya, N., Ogawa, N., Yamada, K. C., Shirotsuki, K., Mikami, I., Hirata, K., Nagano, Y., & Hasegawa, T. (2013). The biological effects of acute psychosocial stress on delay discounting. *Psychoneuroendocrinology*, *38*(10), 2300–2308.

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test' – A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology*, *28*(1–2), 76–81. https://doi.org/10.1159/000119004

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der 'Positive and Negative Affect Schedule' (PANAS). *Diagnostica-Gottingen-*, *42*, 139–156.

Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: A review. *Biological Psychology*, *69*(1), 113–132.

Kurtz-David, V., Persitz, D., Webb, R., & Levy, D. J. (2019). The neural computation of inconsistent choice behavior. *Nature Communications*, *10*(1), 1–14. https://doi.org/10.1038/s41467-019-09343-2

Landolt, K., Maruff, P., Horan, B., Kingsley, M., Kinsella, G., O'Halloran, P. D., Hale, M. W., & Wright, B. J. (2017). Chronic work stress and decreased vagal tone impairs decision making and reaction time in jockeys. *Psychoneuroendocrinology*, *84*, 151–158. https://doi.org/10.1016/j.psyneuen.2017.07.238

Lehrl, S., Triebig, G., & Fischer, B. (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica*, *91*(5), 335–345.

Linden, D. V. D., Keijsers, G. P. J., Eling, P., & Schaijk, R. V. (2005). Work stress and attentional difficulties: An initial study on burnout and cognitive failures. *Work & Stress*, *19*(1), 23–36. https://doi.org/10.1080/02678370500065275

Maier, S. U., Makwana, A. B., & Hare, T. A. (2015). 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.

Margittai, Z., Nave, G., Strombach, T., van Wingerden, M., Schwabe, L., & Kalenscher, T. (2016). Exogenous cortisol causes a shift from deliberative to intuitive thinking. *Psychoneuroendocrinology*, *64*, 131–135. https://doi.org/10.1016/j.psyneuen.2015.11.018

Margittai, Z., Nave, G., Van Wingerden, M., Schnitzler, A., Schwabe, L., & Kalenscher, T. (2018). Combined effects of glucocorticoid and noradrenergic activity on loss aversion. *Neuropsychopharmacology*, *43*(2), 334. https://doi.org/10.1038/npp.2017.75

Margittai, Z., Strombach, T., Joëls, M., Schwabe, L., & Kalenscher, T. (2015). A friend in need: Time-dependent effects of stress on social discounting in men. *Hormones and Behavior*, *73*, 75–82.

Margittai, Z., Van Wingerden, M., Schnitzler, A., Joëls, M., & Kalenscher, T. (2018). Dissociable roles of glucocorticoid and noradrenergic activation on social discounting. *Psychoneuroendocrinology*, *90*, 22–28.

Marin, M.-F., Lord, C., Andrews, J., Juster, R.-P., Sindi, S., Arsenault-Lapierre, G., Fiocco, A. J., & Lupien, S. J. (2011). Chronic stress, cognitive functioning and mental health. *Neurobiology of Learning and Memory*, *96*(4), 583–595. https://doi.org/10.1016/j.nlm.2011.02.016

Miller, R., Stalder, T., Jarczok, M., Almeida, D. M., Badrick, E., Bartels, M., Boomsma, D. I., Coe, C. L., Dekker, M. C. J., Donzella, B., Fischer, J. E., Gunnar, M. R., Kumari, M., Lederbogen, F., Power, C., Ryff, C. D., Subramanian, S. V., Tiemeier, H., Watamura, S. E., & Kirschbaum, C. (2016). The CIRCORT database: Reference ranges and seasonal changes in diurnal salivary cortisol derived from a meta-dataset comprised of 15 field studies. *Psychoneuroendocrinology*, *73*, 16–23. https://doi.org/10.1016/j.psyneuen.2016.07.201

Nater, U. M., & Rohleder, N. (2009). Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: Current state of research. *Psychoneuroendocrinology*, *34*(4), 486–496.

Nitsch, F. J., & Kalenscher, T. (2020). *Keeping a cool head at all times. What determines choice consistency?* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/etyhx

Petrowski, K., Paul, S., Albani, C., & Brähler, E. (2012). Factor structure and psychometric properties of the trier inventory for chronic stress (TICS) in a representative german sample. *BMC Medical Research Methodology*, *12*(1), 42. https://doi.org/10.1186/1471-2288-12-42

Qin, S., Hermans, E. J., van Marle, H. J. F., Luo, J., & Fernández, G. (2009). Acute Psychological Stress Reduces Working Memory-Related Activity in the Dorsolateral Prefrontal Cortex. *Biological Psychiatry*, *66*(1), 25–32. https://doi.org/10.1016/j.biopsych.2009.03.006

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice. *Journal of Economic Literature*, *44*(3), 631–661. https://doi.org/10.1257/jel.44.3.631

Riis-Vestergaard, M. I., van Ast, V., Cornelisse, S., Joëls, M., & Haushofer, J. (2018). The effect of hydrocortisone administration on intertemporal choice. *Psychoneuroendocrinology*, *88*, 173–182.

Rohleder, N., Nater, U. M., Wolf, J. M., Ehlert, U., & Kirschbaum, C. (2004). Psychosocial Stress-Induced Activation of Salivary Alpha-Amylase: An Indicator of Sympathetic Activity? *Annals of the New York Academy of Sciences*, *1032*(1), 258–263. https://doi.org/10.1196/annals.1314.033

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, *5*(17), 61–71.

Schweda, A., Faber, N. S., Crockett, M. J., & Kalenscher, T. (2019). The effects of psychosocial stress on intergroup resource allocation. *Scientific Reports*, *9*(1), 18620. https://doi.org/10.1038/s41598-019-54954-w

Starcke, K., & Brand, M. (2016). Effects of stress on decisions under uncertainty: A meta-analysis. *Psychological Bulletin*, *142*(9), 909–933. https://doi.org/10.1037/bul0000060

Stöber, J. (1999). Die soziale-erwünschtheits-skala-17 (SES-17): Entwicklung und erste befunde zu reliabilität und validität [The social desirability scale-17 (SDS-17): Development and first findings on reliability and validity]. *Diagnostica*, *45*(4), 173–177.

Strobel, A., Beauducel, A., Debener, S., & Brocke, B. (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift Für Differentielle Und Diagnostische Psychologie*.

Sugden, R. (1991). Rational Choice: A Survey of Contributions from Economics and Philosophy. *The Economic Journal*, *101*(407), 751–785. JSTOR. https://doi.org/10.2307/2233854

Tasker, J. G., Di, S., & Malcher-Lopes, R. (2006). Rapid glucocorticoid signaling via membrane-associated receptors. *Endocrinology*, *147*(12), 5549–5556.

Varian, H. R. (1982). The nonparametric approach to demand analysis. *Econometrica: Journal of the Econometric Society*, 945–973. https://doi.org/10.2307/1912771

Varian, H. R. (1991). *Goodness-of-fit for revealed preference tests*. Department of Economics, University of Michigan Ann Arbor.

Vinkers, C. H., Zorn, J. V., Cornelisse, S., Koot, S., Houtepen, L. C., Olivier, B., Verster, J. C., Kahn, R. S., Boks, M. P., & Kalenscher, T. (2013). Time-dependent changes in altruistic punishment following stress. *Psychoneuroendocrinology*, *38*(9), 1467–1475.

Von Dawans, B., Fischbacher, U., Kirschbaum, C., Fehr, E., & Heinrichs, M. (2012). The social dimension of stress reactivity: Acute stress increases prosocial behavior in humans. *Psychological Science*, *23*(6), 651–660.

von Dawans, B., Kirschbaum, C., & Heinrichs, M. (2011). The Trier Social Stress Test for Groups (TSST-G): A new research tool for controlled simultaneous social stress exposure in a group format. *Psychoneuroendocrinology*, *36*(4), 514–522. https://doi.org/10.1016/j.psyneuen.2010.08.004

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Zellner, D. A., Loaiza, S., Gonzalez, Z., Pita, J., Morales, J., Pecora, D., & Wolf, A. (2006). Food selection changes under stress. *Physiology & Behavior*, *87*(4), 789–793.

Table 1

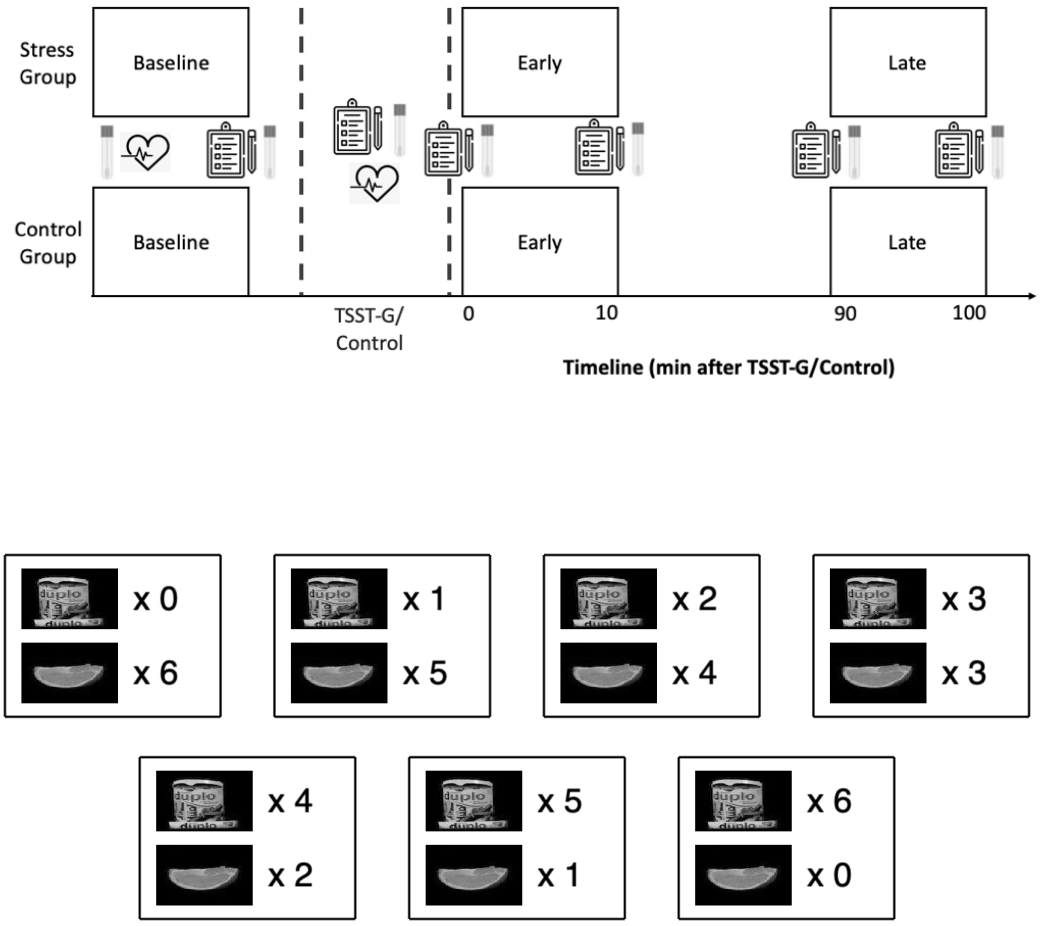
Demographic and trait measures data per experimental group

Demographic Variable	Experimental group	Control group	BF_{10} (for group diff.)
N	N = 75	N = 69	
Age	$M = 24.10, SD = 5.14$	$M = 24.80, SD = 4.48$	0.27
Sex	$female: 39, male: 36$	$female: 37, male: 32$	0.21
University Degree	$yes: 52, no: 23$	$yes: 50, no: 19$	0.19
TICS	$M = 140.00, SD = 27.00$	$M = 149.00, SD = 24.20$	1.11
BFI-10			
Extraversion	$M = 6.33, SD = 1.95$	$M = 6.49, SD = 1.75$	0.20
Agreeableness	$M = 6.65, SD = 1.78$	$M = 6.35, SD = 1.68$	0.30
Conscientiousness	$M = 6.76, SD = 1.63$	$M = 6.52, SD = 1.67$	0.25
Neuroticism	$M = 5.89, SD = 1.98$	$M = 6.06, SD = 1.92$	0.20
Openness	$M = 7.29, SD = 1.89$	$M = 7.71, SD = 1.78$	0.42
BIS/BAS			
Behavioral Inhibition	$M = 15.00, SD = 3.85$	$M = 14.10, SD = 3.61$	0.52
Drive	$M = 7.76, SD = 2.30$	$M = 7.51, SD = 1.84$	0.23
Fun Seeking	$M = 7.64, SD = 1.63$	$M = 7.36, SD = 1.67$	0.29
Reward Responsiveness	$M = 8.20, SD = 2.32$	$M = 7.87, SD = 1.98$	0.26
QDQ	$M = 36.10, SD = 5.78$	$M = 33.80, SD = 5.53$	2.29
SDS	$M = 21.40, SD = 3.00$	$M = 21.00, SD = 2.86$	0.23
MEQ	$M = 15.90, SD = 1.75$	$M = 15.80, SD = 2.04$	0.18
MWT-B	$M = 27.20, SD = 4.18$	$M = 28.20, SD = 4.18$	0.46

Bayes factors were calculated using t-tests for continuous dependent variables and contingency tables for categorical dependent variables with default priors. Abbreviations: TICS = Trier Inventory of Chronic Stress, BFI-10 = 10-item version of the Big Five Inventory, BIS/BAS = Behavioral Inhibition/Activation Scale, QDQ = Quick Delay Questionaire, SDS = Social Desirability Scale, MEQ = Morningness-Eveningness-Scale, MWT-B = multiple-choice vocabulary test (German: *Multipler Wortschatz Test*).

Figure 1

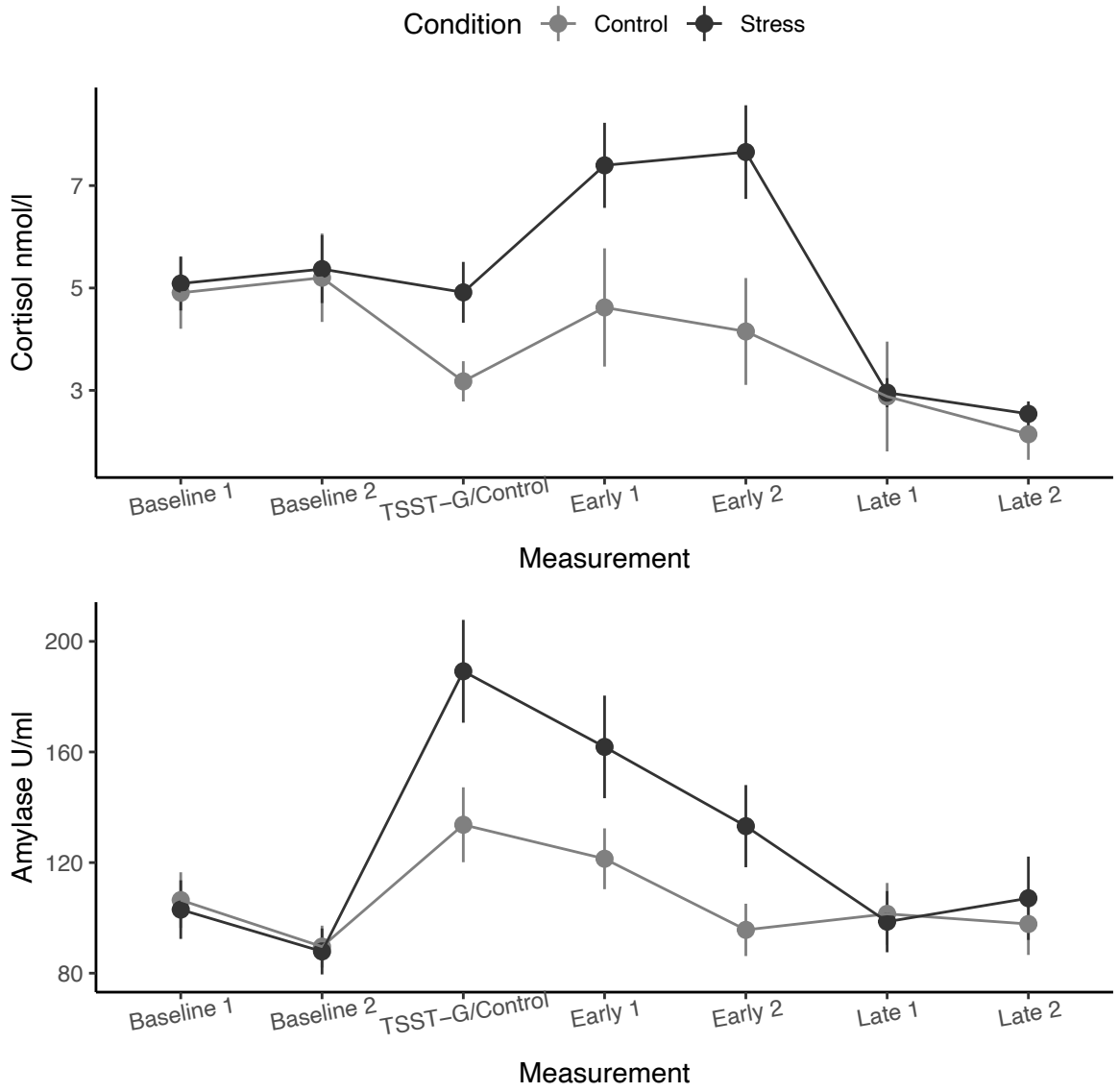
Experimental timeline, measurements and task



Upper panel. Experimental timeline including measurements. Saliva samples were taken before and after each completion of the food choice task as well as during the TSST-G/control procedure. Mean heartrate was measured during baseline and during the TSST-G/control procedure. The self-report measures were taken before and during the TSST-G/control procedure, as well as before and after every subsequent completion of the food choice task. The time on the x-axis is not true to scale. **Lower panel.** Example trial of the food choice task. In this particular trial, the choice set consists of 7 different snack bundles ranging from 6 orange pieces and 0 chocolate bars to 0 orange pieces and 6 chocolate bars. A choice set is always given by all choice bundles whose cost fully exhausts but not overspends the budget given the current prices of both snack types. By increasing and decreasing the prices as well as varying their ratio, we receive the 22 different choice sets of the experiment. For the example trial, chocolate bars and orange pieces were equally expensive, allowing to purchase exactly 6 units of snacks. Importantly, participants were only presented with the final choice set, but not prices or budgets to minimize the difficulty of understanding the task.

Figure 2

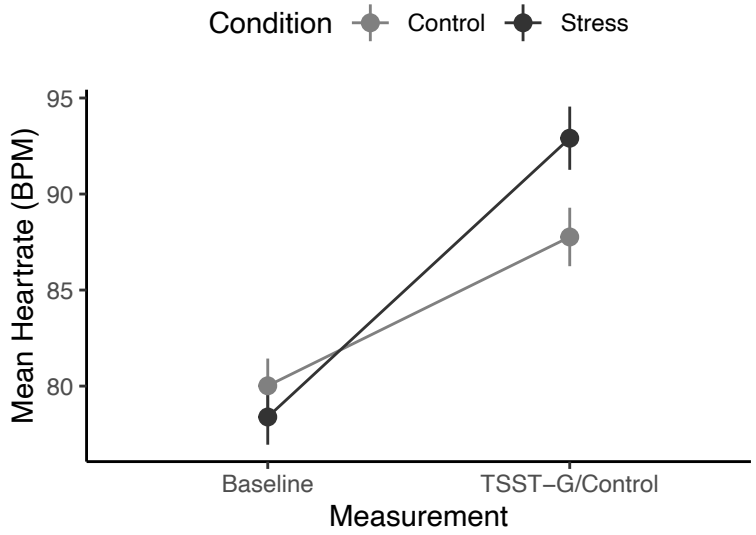
Manipulation check: Cortisol & alpha-amylase levels



Depicted are the mean cortisol (top) and alpha amylase (bottom) levels across the experimental timeline. Error bars show the standard error of the mean. The sample time-points depicted on the x-axis refer to the pre- and post-TSST-G samples, as shown in figure 1. For cortisol, group differences peaked after the first behavioral test and decayed over the course of the experiment. For alpha amylase, group differences peaked during the TSST-G/control procedure and decayed over the course of the experiment.

Figure 3

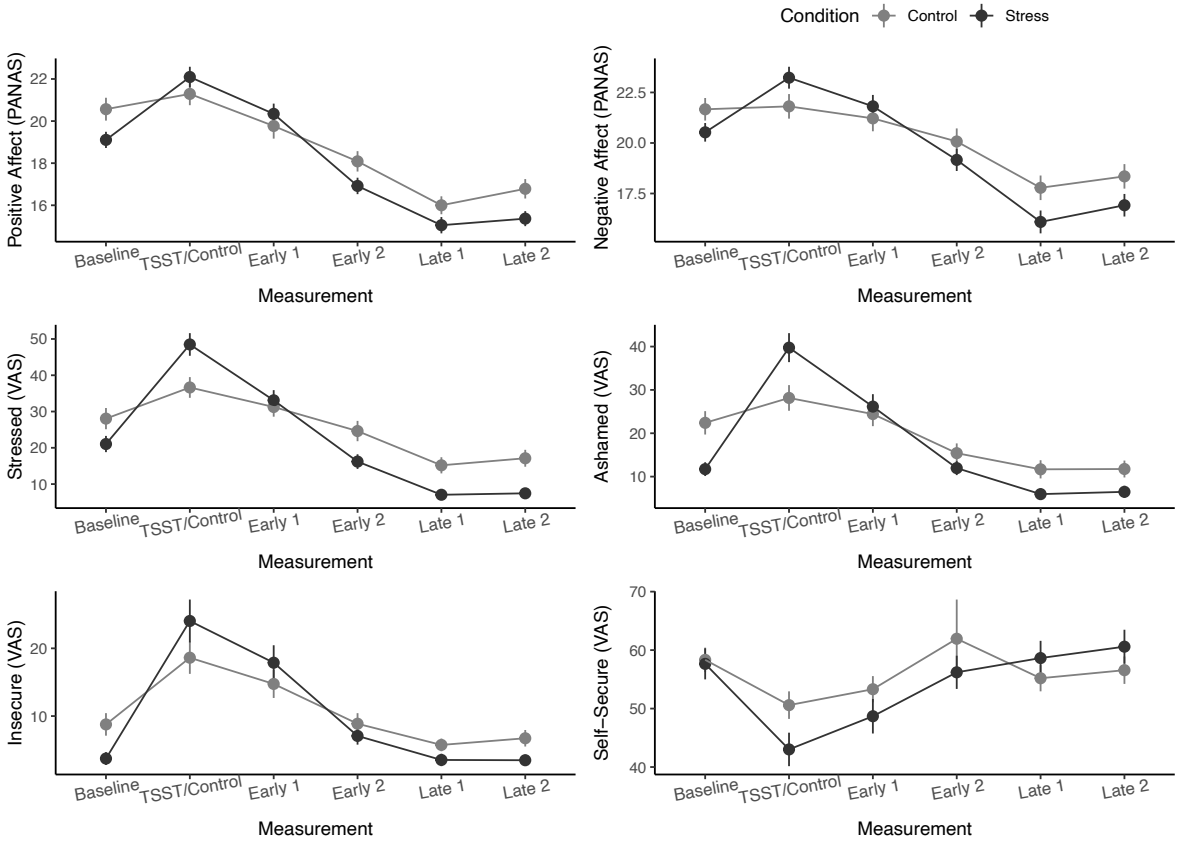
Manipulation check: Mean heartrate



Depicted are the mean heartrate (in beats per minute) at baseline and during the TSST-G/control procedure. Error bars show the standard error of the mean.

Figure 4

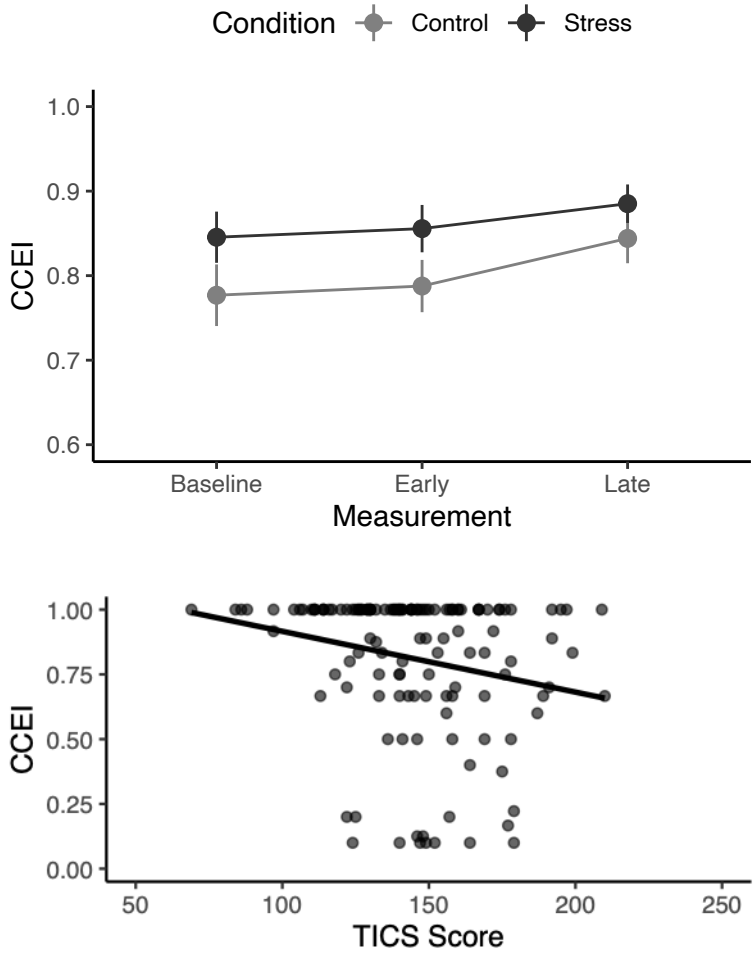
Manipulation check: Self-reported affect



Depicted are the mean response levels for all self-reported affect measures across the experimental timeline. Error bars show the standard error of the mean. Generally, group differences for self-reported affect peaked during the TSST-G/control procedure and decayed or even reversed over the course of the experiment. This reversal might be explained via arousal/fatigue dynamics.

Figure 5

Acute and chronic stress and choice consistency



Upper panel. Depicted are the mean choice consistency levels across the experimental timeline, measured by the Critical Cost Efficiency Index (CCEI). Error bars show the standard error of the mean. Choice consistency was mostly stable over the course of the experiment. Importantly, we found strong evidence against an interaction of time point and group. Note, that while descriptively there appears to be an overall difference between the two groups regarding choice consistency, there is no evidence favoring the inclusion of the experimental group factor compared to the null model ($BF_{10} = 0.76 \pm 0.89\%$). **Lower panel.** Each point in the scatter plot represents the data pair of the TICS score and the CCEI during the baseline measurement of a single subject. The trendline shows a least-squares regression. A negative relationship of chronic stress and choice consistency was almost 10 times more likely than a null relationship given the data ($BF = 9.03$). The empirical correlation coefficient was $\rho = -0.22$, the posterior estimate of the association under the assumption that it is non-zero was $\rho = -0.21$. A visual inspection of the plot reveals that many participants are perfectly consistent. A visual inspection of the scatter plots reveals that many subjects' decisions, in fact, perfectly consistent. This is in line with previous studies (i.e. Choi et al., 2014).

Supplemental Table

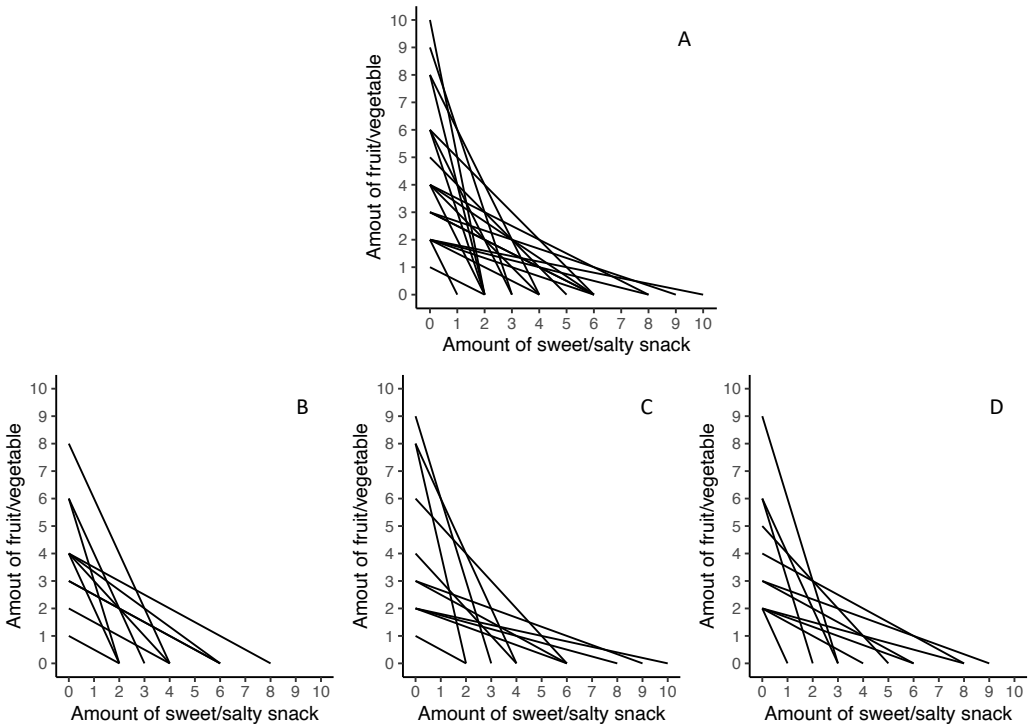
Evidence for the effectiveness of the stress manipulation

Measure	BF_{10}	BF_{incl}
Positive Affect (PANAS)	> 100	> 100
Negative Affect (PANAS)	> 100	> 100
Stressed (VAS)	> 100	> 100
Ashamed (VAS)	> 100	> 100
Insecure (VAS)	> 100	$15.89 \pm 1.50\%$
Self-Secure (VAS)	> 100	$0.24 \pm 3.71\%$

The second column denotes the Bayes factor in favor of the full model (main effects of time point and group and interaction effect) compared to the null model (no main effects, no interaction effect). The third column denotes the Bayes factor in favor of the full model compared to a reduced model (both main effects but no interaction). Generally, a Bayes factor larger than 10 is considered strong evidence, and a Bayes factor larger than 100 extreme evidence.

Supplemental Figure S1

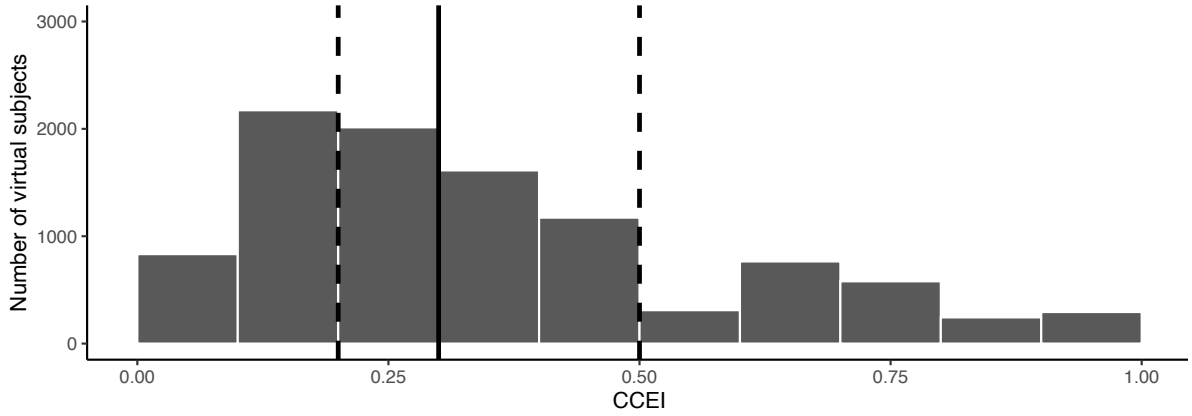
Budgetlines for collection of all possible trials and three sample time points



Panel A shows the budgetlines for the collection of all 22 possible trials. A budgetline shows the all choice bundles whose cost fully exhausts but not overspends the budget given the current prices of both snack types. For example, if the decision maker has a budget of 6€ and both orange pieces and chocolate bars cost 1€ per piece, then the budgetline is given by 7 different snack bundles ranging from 6 orange pieces and 0 chocolate bars to 0 orange pieces and 6 chocolate bars. The slope of the budgetline can be interpreted as the exchange rate between the two different snack types. A higher budget given equal prices allows participants to buy a larger total amount of snacks. This is expressed by the distance of the budgetline to the cartesian origin. Importantly, no choice bundle in any trial contained more than 10 or less than 1 snack units in total. For each participant and time point, 11 trials were randomly sampled from the collection. Panel B, C, and D each show the budgetlines for such a sample of 11 trials.

Supplemental Figure S2

Results of simulation to determine power of design to detect choice inconsistency



To determine the power of our design to detect inconsistent choice behavior we conducted a simulation study with 10.000 virtual subjects. Each virtual subject was assigned 11 random trials from the collection of all 22 possible trials. Virtual subjects chose bundles for each trial randomly with equal probability. Depicted is a histogram of the Critical Cost Efficiency Index (CCEI) for the virtual subjects. The virtual subjects had a median CCEI of $Median = 0.3$ $(IQR = 0.3)$. 97.2% of the virtual subjects showed some degree of inconsistency ($Bronar's\ Power = 0.972$).