Financial Speculations, Stress, and Gender: A Laboratory Experiment[1]

Lubomír Cingl[2] & Miroslav Zajíček[3]

September, 2017

Abstract

In this paper we study the effects of acute stress on speculative behavior using a controlled laboratory experiment with 208 healthy subjects. We employ a recently introduced measure that captures individual speculative behavior, the Speculation Elicitation Task, and an efficient stress-inducing procedure, the Trier Social Stress Test for Groups, and pay special attention to the gender-specific effects. Our design allows for the separation of the main channels behind the treatment effects. We observe strong gender differences: The treatment – stress-inducing – procedure increases men's willingness to speculate compared to control men, but it decreases it for women by about the same amount. As we do not observe any change in the task-specific risk-preferences, and only a little change in the strategic expectations of shift in others' behavior and in the abilities for k-level thinking, we conclude that the behavioral change is driven by the change in preferences, although in the opposite directions for both genders. The analysis of salivary cortisol and subjective mood shows that the subjects were under a considerable level of stress.

[1] This research was supported by a grant from the CERGE-EI Foundation under the program of the Global Development Network RRC16. All opinions expressed are those of the authors and have not been endorsed by CERGE-EI or the GDN. Financial support from research grant GA15-03488S from the Grant Agency of the Czech Republic (GACR) and from grant IGA 6/2016 from the Internal Grant Agency of the University of Economics in Prague are also gratefully acknowledged. The researchers are thankful for the comments received at the GDN Annual 2016 Workshop in Prague, Paul Wachtel and Peter Katuščák.

[2] Department of Institutional, Environmental and Experimental Economics, Faculty of Economics, University of Economics in Prague, Nám. W. Churchilla 4, 130 67 Praha 3, Czech Republic.

[3] University of Economics in Prague, Nám. W. Churchilla 4, 130 67 Praha 3, Czech Republic.

Introduction

Traders in commodity, financial and stock markets daily process enormous volumes of money through which they directly influence the process of price formation in these markets. In the recent 2008-2009 crisis we have seen that the problems in financial sector eventually spilled over to the real sector in the developing as well as in developed countries, affecting lives of most people on the planet. If such market instability is at least partially exaggerated by irrational exuberance and fuelled by imperfections in cognition of market participants, reducing the may improve the wellbeing of many. Even though traders are carefully selected from a large number of applicants, they are still only humans and therefore the limitations of human abilities; particularly the effects of stress which arise from an uncertain, uncontrollable and novel environment (Dickerson & Kemeny, 2004) also apply to them as has been documented in the literature (Abdellaoui, Bleichrodt, & Kammoun, 2013; Coates & Herbert, 2008; List & Haigh, 2010; Lo & Repin, 2002; Oberlechner & Nimgade, 2005).

In this paper we study how acute stress affects individual willingness to speculate since stress has been show to seriously affect individual behavior and decision making (for a review, see e.g. Starcke & Brand, 2012) and the times of market turmoil can generate substantial amount levels of stress in the market participants. A distinctive and robust feature of these markets is the tendency of prices to deviate from underlying fundamental values and form bubbles, which are predominantly fuelled by purely speculative behavior. Such behavior can actually be rationalized using the central idea of the "greater-fool theory" in that some trading parties have beliefs about the existence of other, naive participants willing to pay even higher prices (e.g. Blanchard & Watson, 1982; De Long, Shleifer, Summers, & Waldmann, 1990a), and stress may affect such behavior in several ways. Studying this phenomenon in the laboratory allows us to eliminate other reasons that may contribute to an instability in financial markets and which could potentially be affected by stress, such as herding (Avery & Zemsky, 1998), momentum trading (De Long, Shleifer, Summers, & Waldmann, 1990b; Hong & Stein, 1999), or limits to arbitrage (Shleifer &

Summers, 1990). Moreover, a careful design helps us to uncover the role of several channels of the behavioral change.

We further focus on gender-specific treatment effects as it has recently been proposed that gender of traders may matter in the bubble formation at least in the laboratory environment: specifically that the presence of women may reduce the size of the bubbles (Eckel & Füllbrunn, 2015), even though other papers show little or no support for this finding (Cueva & Rustichini, 2015; Eckel & Füllbrunn, 2017; Holt, Porzio, & Yingze Song, 2017). In economic experiments women are generally found to be more risk-averse, less overconfident and more avoiding competitive situations than men (Croson & Gneezy, 2009; Charness & Gneezy, 2012; Niederle & Vesterlund, 2011). Studies examining the real-world financial performance of men and women show that women tend to be more cautious investors in that they invest more in the risk-free assets, follow less extreme investment strategies, trade less often, but their performance seems not to be different from men's (Hariharan, Chapman, & Domian, 2000; Jianakoplos & Bernasek, 1998; Niessen & Ruenzi, 2015). Moreover, the psychological literature actually shows that the characteristic of response to stress as well as its magnitude is probably different for men and women (Kajantie & Phillips, 2006; Reschke-hern, Okerstrom, Edwards, & Tranel, 2017; Taylor, 2006; Taylor et al., 2000). Because the trading industry is dominated by men, it is of an utmost importance to investigate the potential changes in the behavior of market participants due to the change in gender composition, a policy that many advocate for.

To measure an individual tendency to speculate in a laboratory environment, we employ a recently introduced Speculation Elicitation Task (SET, Janssen, Weitzel, & Füllbrunn, 2015; Moinas & Pouget, 2013) under the stress and control conditions of the protocol Trier Social Stress Test for Groups (TSST-G, Kirschbaum, Pirke, & Hellhammer, 1993; von Dawans, Kirschbaum, & Heinrichs, 2011). Based on the study of Janssen et al. (2015), the SET score neither depends on individual risk-preferences, gender, nor on cognitive abilities – only on one's beliefs about the SET score of other players and so it represents the "greater-fool" principle supposedly underlying the speculations. The stress protocol TSST-G is well-known in the psychology literature and is considered to be one of the most efficient stress-inducing procedures in terms of the cortisol increase (Allen, Kennedy, Cryan, Dinan, & Clarke, 2014; Dickerson & Kemeny, 2004). Subjects in the treatment (stress) group have to go through a public-speaking and mental-arithmetic exercises in front of a two-member panel while being recorded on a camera; the control (no-stress) group faces tasks that are cognitively similar, but with not stressful aspects. We sampled participants' saliva to check the effectiveness of the manipulation on the levels of hormone cortisol, a well known biomarker of stress (Hellhammer, Wüst, & Kudielka, 2009), and elicited mood before and after the TSST-G to capture also the subjective stress effects (Steyer, Schwenkmezger, Notz, & Eid, 1997).[4] To investigate the effects of gender composition on speculative behavior, we set up 10 all-male, 8 all-female and 8 mixed (50 % males and 50 % females) sessions.[5]

We hypothesized that stress could operate through several channels that we are able to separate: (i) the stressed trading parties could directly change their preferences to speculate, (ii) they may suffer from the deterioration of cognitive capacities needed for the "greater-fool" type of thinking or "level-k rationality" resulting in lower "k" under stress (as suggested in Leder, Häusser, & Mojzisch, 2013, 2015), (iii) the market participants could just expect such lower "k" of others without stress affecting their own behavior, or (iv) stress could affect risk preferences that could play a role in the aggregate market stability (Cahlíková & Cingl, 2017; Coates & Herbert, 2008).[6]

To separate the effect of stress on preferences (channel i) from the effect on k-level thinking abilities (channel ii), we elicited beliefs about the decisions of others in both SET tasks with human counterparts: if stress affects the abilities to operate the level-k thinking, subjective beliefs should then provide the same picture as the decisions made in the SET. To separate channel (iii), the change in the subjective

---

[4] We note that cortisol is a marker of endured stress, but is not the only one and narrowing the effects of stress to the sole effect of cortisol could be misleading; see Everly & Lating (2013) for details about physiological effects of stress.

[5] For each gender composition, one half of the sessions was the treatment (stress) and one half control (no-stress) condition.

[6] In the domain of measuring the effect of acute stress on risk-preferences, different elicitation mechanisms and timings of measurements yield different results, ranging from increased risk-aversion (Cahlíková & Cingl, 2017) to increased risk-seeking (Buckert et al., 2014), while other authors even claim the time-dependency of this effect (Pabst et al., 2013).

beliefs about the stress-effects on other subjects, we let subjects decide in the SET first when they knew that the two other players in the SET (the "counterparts") were from the same room and thus are undergoing the same (stress or a control) procedure, and another time subjects performed in the SET knowing the two other players in the task come from another group of participants that was not stressed.[7] Channel (iv) can be identified from the third decision in the SET framework which was a risk-preference elicitation task. Here subjects were informed that this time it is a computer and not two other people in the SET task who is assigned to the SET with them; and they were informed about the precise probabilities that determined the decision making of the computer.

Apart from being the first to experimentally study the above mentioned questions, we generally contribute to the literature on the effects of acute stress on human economic decision making, a relatively scarce literature that focused on the effects of stress on risk-preferences (see the most recent comparison of studies in Sokol-hessner, Raio, Gottesman, Lackovic, & Phelps (2016)), time preferences (Haushofer et al., 2013; Riis-vestergaard, Ast, Cornelisse, Joëls, & Haushofer, 2017), social preferences (Vinkers et al., 2013; von Dawans, Fischbacher, Kirschbaum, Fehr, & Heinrichs, 2012) willingness to compete (Buser, Dreber, & Mollerstrom, 2017; Cahlikova, Cingl, & Levely, 2017; Zhong, Shalev, Koh, Ebstein, & Chew, 2017), and competitive confidence (Goette, Bendahan, Thoresen, Hollis, & Sandi, 2015).

The SET score has been shown to predict the size of the bubble in experimental asset markets (Janssen et al., 2015) which connects our contribution with this rich behavioral-finance literature (Palan, 2013; Powell & Shestakova, 2016; Smith, Suchanek, & Williams, 1988) We are not aware of another published study focusing on the effects of stress on financial decision making. To the best of our knowledge, only several studies examined the relationship of specific hormones and the financial markets (Cueva et al., 2015; Kandasamy et al., 2014; Nadler, Jiao, Johnson, Alexander, & Zak, 2017), and the effect of the

---

[7] These counterparts that did not undergo the TSST-G procedure participated in the SET task as well and were specifically referred to as "the participants of a different experiment that continued in the laboratory with different tasks on the computers" while any direct mention of stress was avoided. The number of subjects and the gender composition in these groups was the same as in the TSST-G groups.

depletion of concentration resources on the size of bubbles in experimental asset markets (Kocher, Lucks, & Schindler, 2015).

We are aware that traders are a highly selected, self-selected and hard-to-reach group. In order to increase the external validity of our findings, we focus on a key characteristic that can potentially directs the effects of stress on the individual response in SET, the trait anxiety (STAI-T, Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). In a recent study, trait anxiety directed the change in self-confidence under acute stress in that the low-anxious subjects under stress reacted with an increased confidence under stress compared to control and vice versa (Goette et al., 2015). We particularly hypothesized that the participants low in STAI-T could have an enhanced confidence under stress and also increase the SET score, similarly as the real traders.[8]

Interestingly, our main results show that with no regard to the gender specific behavior, there is no effect of stress on the SET score. However, when we look at the men and women separately, the picture changes dramatically. Compared to control, stress causes men to increase their speculative behavior by about the same amount as it causes women to decrease theirs. The investigated channel behind this effect seems to be predominantly the physical stress, and only very little the strategic expectations about the effect of stress on others and the level-k thinking abilities. Remarkably, we do not observe any difference in the risk-task, both between stress and control participants as well as between men and women in both treatments. Women do not change their level of speculative behavior due to treatment intervention in the mixed sessions compared to the all-female sessions, where we observe a sharp decrease. The samples of subjects' saliva revealed that the levels of cortisol in the stress condition were substantially elevated (by 112% on average) while in the control condition they mildly decreased (by 15%). The subjective measure of stress, the mood questionnaire, reveals that the treated subjects felt worse, but not more tired or nervous than the control.

[8] Our unpublished results show that traders from a large Czech energy trading company have by about 0.5-0.6 SD lower STAI-T than a student sample.

Experimental Design

Game: Speculation Elicitation Task

We closely follow the setting of the SET in Janssen et al. (2015) which is based on the Bubble game (Moinas & Pouget, 2013). There is a sequence of three traders who can either accept or reject to buy an asset which has a fundamental value of zero. If a trader accepts to buy the asset, it involves an investment of 3 ECU (experimental currency unit). Then, if the trader succeeds in selling it to the trader who is next in the sequence, she earns 10 ECU. A big caveat is that if the trader is stuck with the asset, either because the next trader does not want to buy it or there is no more trader in the sequence, she gets 0 ECU, its fundamental value. The game-tree with the resulting payoffs is depicted in Figure 1.

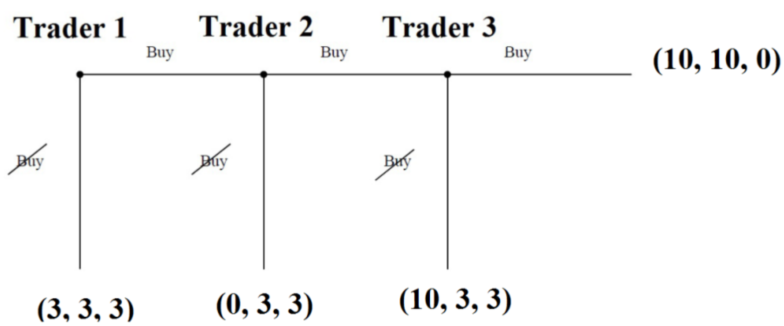


Figure 1: The game-tree of the bubble game. ***Source: Janssen et al. (2015)***

The Nash equilibrium of this game similarly as in the centipede game can be obtained by backward induction: no trade resulting in no bubbles. The crucial feature of the design is that the traders do not know their position in the sequence; rather they receive an imprecise signal from a known distribution. This signal is provided in the form of a "price" of the asset. The price is at the beginning of the sequence drawn randomly from the distribution of $P_1 \in \{10^0, 10^1, 10^2, 10^3, 10^4\}$ and if a trader sells it further, it increases by a factor of 10. Therefore, the price of 1, 10 and 100 is fully informative of not being the last in the sequence while the price of 1.000.000 implies being the last in the sequence for sure. All participants make all decisions simultaneously and if anyone decides to buy, it is assumed that she wants to further sell. The participants are making decisions in the strategy form for all possible prices that can arise and are informed of the implied probabilities of being the last in the sequence, given this price, starting at the highest possible price and then in steps descending to zero. If a price is accepted, it is assumed all lower prices would be accepted as well. The SET score is directly related to the switching point from which on the participant decides to buy as the rank of the switching price P^S, see Table 1.

---Table 1 about here --

The final payoff for a subject is calculated such that first, one price is drawn randomly from the announced distribution, then the position of a trader was randomly determined, while the remaining two traders were chosen from the respective room and their decisions determined subject's payoff.

The participants are then asked in an incentivized way[9] to state their beliefs about the SET scores of other participants by asking the following for each value of the price: "*How many of participants in the today's session start buying at this point?*" Instructions were given first aloud by the experimenter and understanding was checked using a series of detailed on-screen questions.

Treatment: Stress & TSST-G procedure

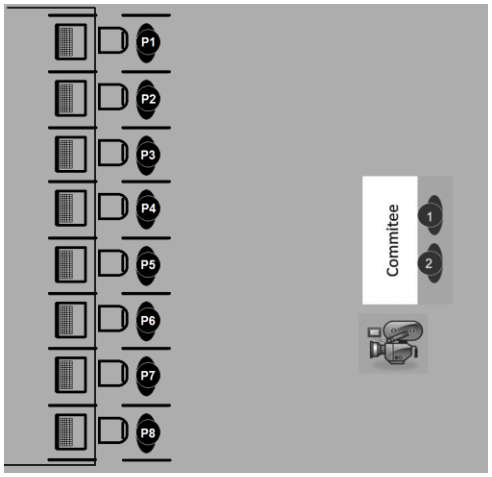
Stress is an instinctive reaction to a perceived threat to a desired goal of an organism; the goal being e.g. the preservation of biological or social self (Kemeny, 2003). It activates two main physiological pathways to restore homeostasis: the sympatho-adreno-medullary (SAM) system and hypothalamus-pituitary-adrenal (HPA) axis. The activation of SAM consists of an instant increase in catecholamines (epinephrine, norepinephrine, dopamine) which stimulate the cardio-vascular system and can be traced by measuring heart-rate, blood pressure or body temperature (Dickerson & Kemeny, 2004). The activation of HPA axis follows in the matter of minutes and as the final product it releases cortisol into the bloodstream. Therefore it has been commonly used as a biomarker of stress, usually combined with a measure of SAM (e.g. heart-rate) to provide evidence of a real-time dynamics of the whole stress reaction (Hellhammer et al., 2009) and its sampling from subjects' saliva is relatively non-invasive. Generally

[9] See the full instructions in the Appendix how the incentives were specified.

speaking, the fast-processed activation of SAM system and HPA axis manifest themselves in rapid focused attention, hypervigilance and promote reflex-like behavior at the expense of goal-directed behavior (Schwabe, Tegenthoff, Höffken, & Wolf, 2010).

Since different stressors trigger different stress reactions, we study the effect of psycho-social stress as it most accurately resembles the conditions a trader faces – a failure represents a threat to a goal of achieving success in their career; and to preservation of their social status derived from the evaluation by other people (Dickerson & Kemeny, 2004). This can be reliably induced in laboratory conditions by the means of the standard psychological protocol Trier Social Stress Test (TSST, Kirschbaum et al., 1993) in its modification for groups (TSST-G, von Dawans et al., 2011). The procedure which is carried out in a different room than the laboratory consists of two parts: public speaking and mental arithmetic. The framing of these was in our setting slightly modified from the original: in the first part subjects were asked to provide a mock job-interview to a job of their dreams in two minutes; and in the second part the subjects were asked to say out loud every second letter in the alphabet in the backward order (e.g. Z, X, V, ...) for one minute. Every subject received a different starting letter. During this process they were standing separated by cardboard wall in front of a panel of two people (refer to Figure 2) who wore white laboratory coats, made notes, and were specially trained not to give any feedback on the subjects' performance. Subjects' performance was recorded by a video-camera operated by one of the panel members. To avoid hearing others, subjects had headphones with an ambient noise on when they were not talking to the committee. The sole procedure took about 30 minutes. At the end of the experiment a proper feedback was carefully administered to the treatment group.

Figure 2: Scheme of the TSST-G room. Source: Adapted from von Dawans et al., (2011)**.**



An important part of the protocol is the full control procedure, where subjects perform cognitively comparable activities, but without the stressful aspects: the speaking part consists of reading a text from a popular magazine in a low voice and in the arithmetic part subjects simply say the alphabet out loud for one minute. The timing of the tasks is the same as in the stress treatment.

To make sure that the stress reaction was induced we sampled subjects' saliva three times (before they receive instructions to the oral part, right after they finish the first oral part, and when they arrived back into the laboratory to ensure the dynamics of the stress reaction is comparable to similar studies) and had it analyzed for the concentration of free cortisol. Apart from that, we measured their mood using a questionnaire that has two parts with a set of 15 questions in each. The outcome is a score for three mood dimensions: good/bad, awake/tired and calm/nervous (MDMQ/MDBF, Steyer, Schwenkmezger, Notz, & Eid, 1997). To check if the attention resources were not differentially depleted, we administered a standard D2 attention test right after the arrival of subjects back to laboratory (Brickenkamp & Zillmer, 1998).

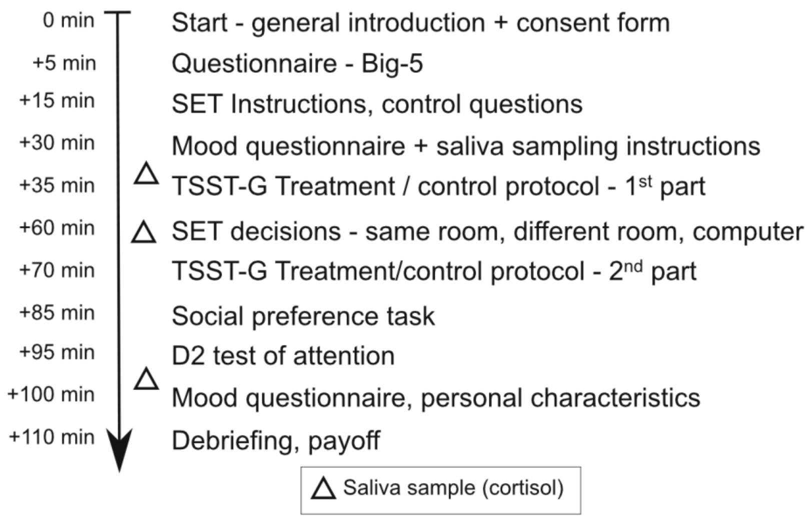
Procedures

First, the invitation email that was sent out using ORSEE (Greiner, 2004), but any mention of stress was avoided in order not to cause a self-selection of subjects. Next, several days prior their participation, subjects received an internet survey with the STAI-T, general risk-preferences (Dohmen et al., 2011), ambiguity preferences (Cavatorta & Schröder, 2014) and other characteristics including questions assessing the physical and mental health status of the subjects: subjects that indicated any problems were not allowed into the experiment.[10] For the day of the experiment, they were informed not to smoke, perform heavy exercise, drink more than three cups of coffee and completely avoid alcohol so that the cortisol level could be accurately measured in their saliva, which was then checked using a short screening questionnaire before they entered the lab. After arrival to the lab subjects, simultaneously with the counterparts, were randomly seated at the computers, general instructions were read aloud and signed an informed consent form (Figure 3 provides a timeline). The counterparts were referred to as the participants of a different experiment, who would take part in some parts of the experiment with the subjects and were seated in a separate room of the laboratory separable by a door, therefore their presence was trustworthy and the fact they did not undergo a stress procedure as well, as confirmed by a question in the final questionnaire. The gender composition was always the same in both groups. Then subjects and the counterparts filled in an on-screen short personality questionnaire (Big-five dimensions of personality, Costa & McCrae, 1992). Then both groups together were read the instructions for the SET by the experimenter and answered a series of detailed on-screen control questions that did not allow a participant proceed without full understanding. The instructions were handed out on paper, together with the tables informing about the probabilities in the SET, so that the participants could look at them when they needed to. Next, they filled in the first part of a MDM questionnaire evaluating their current mood (Steyer et al., 1997). After this, the door separating the room with the counterparts was closed and the counterparts continued with a different task. Subjects were then instructed how to use plastic tubes for saliva sampling and give the first sample – the baseline. Next the headphones with ambient noise were set and when everybody was ready, subjects received instructions for the first part of the TSST-G procedure and had 3 minutes for preparation. The TSST-G procedure in the adjacent room took about 45 minutes in total,

[10] One response from the internet questionnaire was due to technical problems not recorded.

including the behavioral decisions. To exploit the physiological stress response in its full immediate effect, subjects in the TSST-G room had laptops prepared at a desk behind them (Figure 2), where they would sit and make decisions right after the speaking part. The SET was done simultaneously by all subjects after the last subject finished the first oral task, the job-interview, and while working on the computer, the second sample of saliva was given. The SET was administered three times: (i) subjects were informed the other two traders in the sequence would be two other participants from the same room as subjects; (ii) next they informed the other two traders were "participants of the other experiment that stayed in the lab and proceeded with a different task at computers"; and (iii) they were informed the other two traders would in fact be computer making decision according to probabilities known to subjects. Subjects could change their mind once in each of the three decisions. The order of the first two decisions was counterbalanced across sessions while the decision against computer was always the last. The first two decisions were accompanied by an incentivized elicitation of beliefs. After performing in the second oral task, the mental arithmetic part of TSST-G, a simple social-preference game was administered, which is not reported on in this paper. When the TSST-G procedure ended, subjects returned to the lab, filled in the D2 attention test, the second part of the on-screen mood questionnaire, a brief questionnaire on their personal characteristics and were paid in private. They gave the third saliva sample when they started working at the computer, that is right after the D2 attention test. Subjects as well as the counterparts had been informed that for their payoffs, one decision at random would be chosen from the two games they had performed in; two decisions in total, and all expectation-decisions would be paid for unconditionally like the show-up fee. The treatment group received a careful debriefing before the payoffs, when the counterparts had left the laboratory. The experiment was conducted in the Czech language, the head of the TSST-G committee was a male with a female colleague, the experimenter was male and the secondary experimenter who managed the session with the counterparts was female.

Figure 3: The timeline of the experiment



Results

Sample

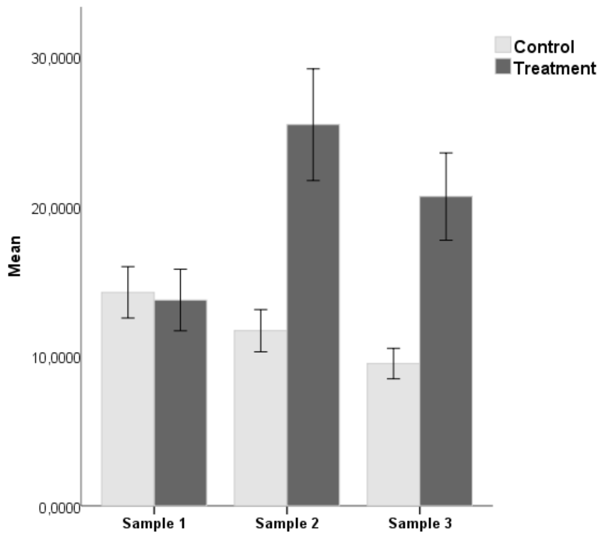
The 26 experimental sessions were executed in two batches with identical procedures in May 2016 and May 2017 with 208 subjects, with additional 208 subjects as their counterparts. Recruitment was done using ORSEE (Greiner, 2004) and the experimental tasks were programmed in Z-TREE (Fischbacher, 2007). We invited 8 subjects per session, and executed either the stress-inducing or the control procedure per session. There were 10 sessions with only male subjects, 8 sessions with 50% of male and female subjects, and 8 session with only female subjects. All sessions started either at 15:00 or 17:30 in order to avoid any problems with cortisol measurement which fluctuates over the course of the day (Nicolson, 2007). Each day, one control session and one treatment session were executed in a counter-balanced order. The average length of a session was a little less than two hours and the average payment was 505 CZK (around 19 EUR). The conversion rate was 1 ECU = 25 CZK.

Randomization check reveals no significant differences between the treatment and control groups in their baseline attitudes to risk and ambiguity, trait anxiety and personality dimensions that were measured in the online questionnaire sent several days prior to the day of participation (F=0.99, only males F=1.52; only Females F=1.72, see Table 8).

Stress response - cortisol

We assess the efficiency of the treatment manipulation through the elevation of salivary cortisol level in the treatment vs. the control group. Subjects sampled their saliva first before they received the instructions to TSST-G (Sample 1) to check whether the groups did not differ before the TSST-G procedure. Sample 2 was collected after the public speaking part of the TSST-G and Sample 3 was collected right after they had returned back to the laboratory. The results are presented in Figure 4. The baseline levels are not different from each other (t-test p=0.70) while the treatment and control groups differ in the second and the third samples significantly (both p<0.001). Since baseline cortisol could be highly individual (Buser et al., 2017), we also compare only the percentage change between Sample 2 and Sample 1 (baseline), and between Sample 3 and Sample 1. When first looking at the change between Sample 2 and to baseline, the cortisol level in the treatment group increased by 112 %, while it decreased by 15 % in the control group. Because we are interested in the gender differences in behavior, we also check the effectiveness of manipulation for both genders apart. For men, the increase was a little steeper as their cortisol rose by 120 %, while for women the increase was about 101 %, but the difference was is not significant (t-test p=0.54). Comparing Sample 3 to baseline yields an increase of 71 % in the treatment and a decrease of 27 % in the control group. When we break it down to gender we observe that in women cortisol level was more stable elevated than in men as they show an increase of 82 % while men only of 60 %, but this difference is again not significantly different from zero (t-test p=0.28). Overall, this provides evidence that the subjects in the treatment group were on average under considerable levels of physiological stress while the subjects in the control group were not.

Figure 4: Cortisol levels



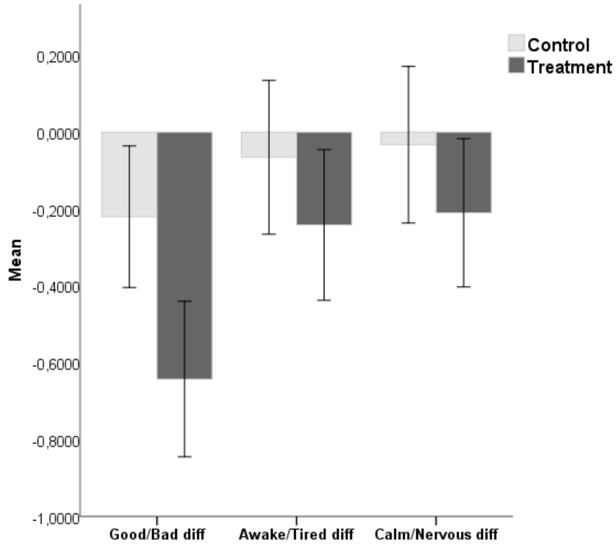
Notes: Darker color indicates the treatment group. The Y axis scale units are nmol/l. Error bars show 95% confidence intervals. Sample 1 was taken before the TSST-G procedure, sample 2 in the course of the TSST-G procedure and Sample 3 after the return to the laboratory from the TSST-G room.

Stress response - Mood

A subjective nature of the TSST-G protocol was measured using the MDM questionnaires. [11] Figure 5 shows the score changes in the three dimensions calculated as the score measured *after* minus *before* the TSST-G protocol. Overall, subjects felt by about 0.42 SD worse (Wilcoxon rank-sum test, p<0.01), but no differences were found in the awake/tired dimension (p=0.24) or the calm/nervous dimension (p=0.19). When we look at the mood responses by gender, men seem to react in the good/bad and the calm/nervous dimensions, but not in the awake/tired one (p<0.01, p=0.06 and p=0.53, respectively). Interestingly, we do not find the mood differences for women in either of the dimensions (good/bad p=0.27, awake/tired p=0.3, calm/nervous p=0.94), which suggests that for women, the psychological reaction is smaller than for men.

[11] In session 1 where the control manipulation of TSST-G was administered, the second part of MDM questionnaire was due to technical problems not administered. The presented results thus slightly overstate the real effect.

Figure 5: Differences in mood.

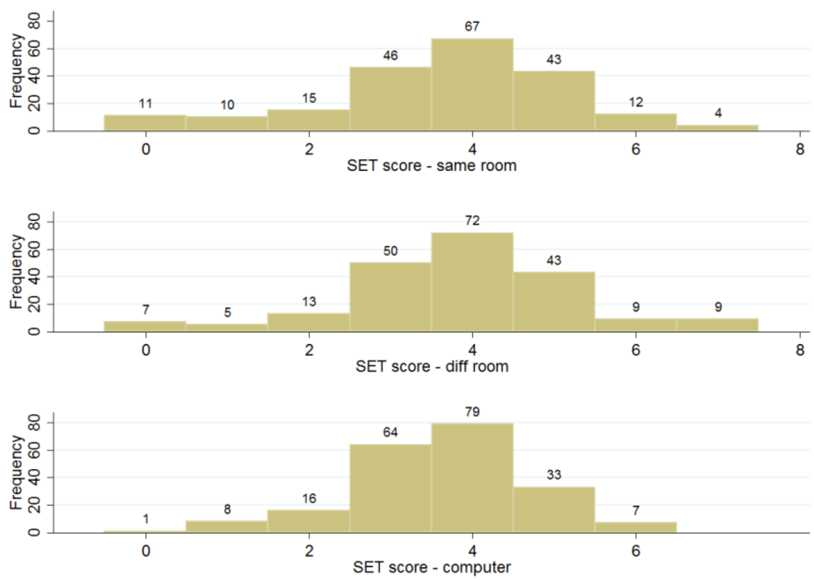


Notes: Darker color indicates the treatment group. Error bars show 95% confidence intervals. Mood change is measured as the difference between the standardized score in the respective dimension before the TSST-G procedure and after the procedure. Y-axis scale units are the standard deviations from the mean.

Speculations: SET scores

The SET score ranges from 0 to 7; a higher SET score implies that subjects speculate more in that they want to sell the worthless asset to a greater fool who is however less likely to appear. First we examine the SET score for all three decisions, when the treatment and control are pooled together: (i) against other two traders from the same room (*SR_SET*), (ii) against other two from the other room (*DR_SET*), and (iii) against computer (*COMP_SET*). We observe in Figure 6 that there are no large differences in the distribution of the SET scores when the other traders in the sequence are from the same and from the different rooms. Note that several subjects with SET=7 (4 in the same-room and 9 in the different-room setting) decided to buy even at the highest price which implied they would not be able to resell the asset with 100 % chance, which casts doubts on their understanding of the game. Interestingly, none had SET=7 in the risk-task, which suggests that subjects who set SET=7 in the two preceding tasks may either have had pro-social motivations to accept the loss of 3 ECU when it would lead to the gain of 10 ECU of another participant, or that there was a learning effect despite having no feedback on their decisions.

Figure 6: SET scores by decision.

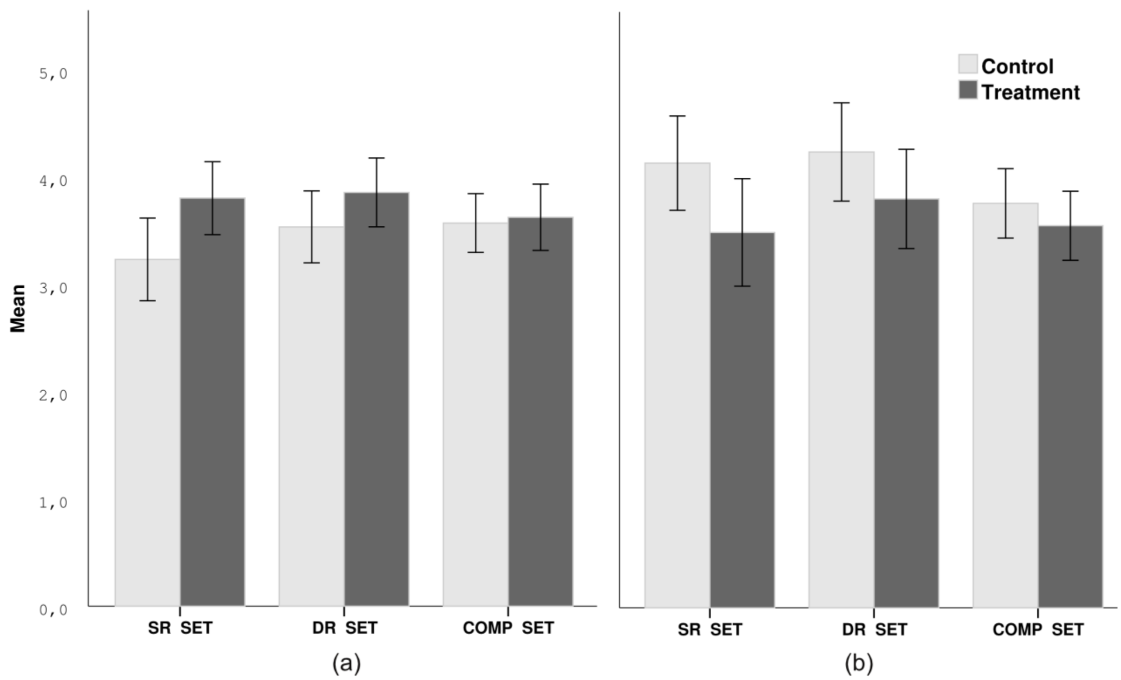


Next we turn our attention to the effects of treatment on the pooled sample, and to the gender specific treatment effects. We present histograms of the decisions in Figure 8 in the Appendix and the summary statistics in Table 2, first lines of panels A, B and C. In terms of the general differences between the treatment and control groups in the three decisions, none comes out significant: the *SR_SET* (Wilcoxon rank-sum test, p=0.98), the *DR_SET* (p=0.63) nor the *COMP_SET* (p=0.54).

---insert Table 2 around here ----

Now we focus on the gender specific reactions that are summarized in Panel A of Table 2 and for better clarity also in Figure 9. Starting with the SET when other two traders were from the same group (SET_SR) we observe a strong gender difference in the reaction to stress manipulation: while score of the males in the stress group is by about 0.57 points larger compared to control (Wilcoxon rank-sum test, p=0.035), women in the stress condition scored by 0.65 (p=0.045) less than women in the control. Women in the control have by 0.9 higher score than men (p<0.001) while in the treatment group we observe a difference of 0.3 point that is not statistically significant (p=0.46). These results are confirmed in column 2 of the regression analysis presented in Table 3 where we regress the *SR_SET* score on the dummy indicating assignment to the stress treatment, dummy indicating whether the subject is a female, and their interaction. The interaction term shows that the size of the difference-in-differences effect is about 1.2 points on the 0-7 scale (p=0.014), which can be considered sizeable.

Figure 7: Treatment effects by gender.



Note: Darker color indicates the treatment group. Panel A shows mean SET scores of males and panel B of females. Error bars show 95% confidence intervals.

Another decision in the SET context was done when the other two traders were from a different room (*DR_SET*) who clearly were not stressed. A different behavior of the treatment group in this task compared to *SR_SET* could uncover the strategic expectations about the stress-effects on the behavior of others. To start with, however, the difference in dealing with traders from the same and from the different room could be driven by a pure out-group effect (see e.g. Charness, Rigotti, & Rustichini, 2007; Chen & Li, 2009; Sutter, 2009), which should be captured by the size of the difference between the *SR_SET* and *DR_SET*[12] in the control group only, and anything on the top of that should reflect the strategic expectations of the stress effects. When we compare the scores of *DR_SET* and *SR_SET* (Table 2, panels A and B, columns 2 and 3, males and females only), we observe a little bit higher scores for *DR_SET* in all gender-specific decisions, but with the exception of males in the control group (paired t-test p=0.06), all come out insignificantly different from zero. The effect of dealing the out-group compared to in-group is thus negligible.

---

[12] Note that the order of these two decisions was counterbalanced across sessions.

A glance into Table 2 tells us that the treatment effects in *DR_SET* are qualitatively similar as in case of *SR_SET*, but smaller in magnitude: men in the treatment group scored by 0.32 points higher than men in the control (p=0.2, compare to 0.57 in *SR_SET*) and treated women scored by 0.44 less than control women (p=0.06, compare to -0.65 in *SR_SET*). The interaction term *Female X Treatment* in *DR_SET* (-0.759, column 2 in Table 3) is about 62 % of the one in the *SR_SET* (-1.217, column 2 in Table 4). Overall, the treatment effects in this task, i.e. dealing with a group that is clearly not-stressed, were not principally different compared to the situation when dealing with the participants from the same group, just smaller. The comparison of *SR_SET* and *DR_SET* thus generally reveals that the strategic expectations about the effects of stress on others seem to play a minor role in the explanation of the speculative behavior, and it is not due to the pure out-group effect.

---insert Table 3 and 4 around here ----

Speculative behavior could change under stress due to the change in risk-preferences. The third decision in the SET framework (*COMP_SET*) serves us as the task-specific risk-preference measure. The variable *COMP_SET* is not statistically different across the treatment status, nor gender (panel C in Table 2) and the interaction of the two (Table 5, column 2). The insignificance of the coefficient on the Female dummy is interesting since it means that in the SET-specific setting we do not observe any gender differences even in the control group, while we do find that men are more risk-seeking than women in the survey question in the internet questionnaire (Wilcoxon rank-sum, p=0.04). The way the risk-attitudes are measured thus seems to play a role (Filippin & Crosetto, 2016). A potentially-gender-specific change in risk-preferences as a channel of the stress-induced change in speculative behavior can thus be ruled out.

---insert Table 5 around here ----

Since in Goette et al. (2015) it is argued that trait anxiety modulates relative confidence which may also play a role in the SET score, in columns 3 and 4 of Table 3 and Table 4 we add standardized[13] trait anxiety score (STAI)[14] and its interaction with treatment,. Neither of the coefficients of variables *STAI* or *STAI X Treatment* is significant and the coefficients of the main variables do not change in their magnitude nor in significance. We further perform a robustness check in the Table 9 and

---

[13] We find a significant level difference in the STAI-T variable between genders, therefore we standardize the variable in terms of standard deviation change from the mean score in the gender-specific group.

[14] The original protocol has two parts: the short-term, state anxiety and the long-term, trait anxiety. In this paper the STAI refers solely to the long-term trait anxiety while the state anxiety was not administered.

Table 10 in the Appendix where we split the sample into low-STAI and high-STAI groups according to median. Coefficients in the regressions run only on the low-anxiety group (columns 2, 5 and 8) are larger in magnitude than from the whole sample and keep the significance levels, while the results from the high-anxiety group (columns 3,6 and 9) are all smaller in magnitude and all insignificantly different from zero. This suggests that the results are driven by the low-anxious subgroup. From our own measurements of the trait anxiety carried out on 16 professional traders of a large energy company and 8 final applicants for this position comes out that they have by 0.5 and 0.6 SD smaller STAI scores than a student sample, respectively (Cingl & Zajíček, unpublished results). This implies that traders' behavior should be affected by acute psychosocial stress, provided the assumed link is true.

Then, we analyze the average beliefs of the participants,[15] again once in the decision when other traders were from the same room (variable *beliefs_SR*) and when they were from the different room (*beliefs_DR*). Panels D and E in Table 2 show that without considering gender, the difference between treatment and control in both beliefs variables is negligible. When broken down to gender-specific effects, the sign of the difference between the treatment and control is the same as in the real decisions, but they are much smaller in magnitude and insignificant. The regression analyses where the beliefs variables are the dependent variables reveal that only the coefficient of the interaction term *Female X Treatment* is significant only in some specifications, only for the "same-room" decision (Table 6 and Table 7 in the Appendix), and is of about half the size of the respective coefficient in the *SR_SET* regression. If we assume that the beliefs mirror the abilities of the level-k thinking, our results suggest they were only marginally affected and thus seem not to be the driving channel of the behavioral change.

Is the change in behavior driven by the physical reaction represented by the cortisol change? Table 11 in the Appendix presents the OLS estimation of the correlations between *SR_SET*[16] with the percentage cortisol increase and the average treatment effect on the treated (ATT) using the instrumental variable (IV) estimation when the cortisol increase is instrumented by the dummy Treatment. We observe that while the plain correlations are insignificant, the IV estimates are significant and similar in sign and magnitude as the main results in Table 3. This suggests that while in general the change in cortisol does not relate to the speculative behavior, the change due to the Treatment intervention does replicate the original Intention-to-treat results well. Physical reaction to stress due to the stress treatment is the plausible cause of the behavioral effects.

---

[15] These were calculated as the average believed SET score of others elicited after the main decision in an incentivized way.

[16] Results for the *DR_SET* in this and the next sections available upon request.

The stress-inducing treatment procedure might also exhaust subjects' concentration resources and the effects could thus be driven by the differences in attention. Attention was measured with the number of errors in the standard D2 attention test at the end of the experiment. We observe that the treatment group produced slightly more errors[17] than the control group (M_control=3.9, M_treatment=5.6, rank-sum p=0.007), which persists even after the exclusion of several obvious outliers (p=0.07). Therefore we repeat the exercise with the correlations and the ATT effects from the previous paragraph to identify which of the variation could be attributed to the Treatment intervention. Table 12 presents the results that indicate very weak relationship between the number of errors and the *SR_SET* score (column 3), but the IV regressions show completely no significant relationship. The change in attention as the channel of the behavioral change can thus be also ruled out.

Lastly we look at how the presence of the other gender in the same session influenced the results. There were single-sex (all-male and all-female) or mixed sessions where strictly 50 % men and women were recruited, both to the "same-room" as well as into "different-room" groups. Examining the gender-specific behavior in *SR_SET* in the mixed sessions leads us to note that we get to rather small numbers in the compared cells, as there were 16 men (women) in treatment and 16 men (women) in control group in the mixed sessions. Males in the mixed control sessions speculate less than in all-male control sessions, but the difference between the behavior of males in treatment is not different between mixed and all-male sessions. As the interaction term is insignificant (col. 3, Table 13 in the Appendix), the size of the difference between the treatment and control group is not different in the mixed and all-male sessions (p=0.211). This suggests that men, in the presence of women, react under treatment only a little more than when they are not around. For females, the difference between treatment and control in the all-female sessions is about -1 point (col. 6, Table 13), but this difference gets much smaller and the point estimate is even positive in the mixed sessions, as shown by the positive interaction term (linear combination of *Treatment + Treatment X Mixed*: beta=0.19, p=0.695). For women, the presence of men in the session results in diminished treatment effects, as they do not decrease the speculative behavior as women in the all-female group. Extrapolating from our results to the real-world situation in the trading floors where almost no women are present, having roughly equal proportion of men and women would likely decrease speculative tendencies of men in the quiet times, but in case of stress in the markets, they would increase their speculative behavior to the same level as if no women were present. For women, there would likely be no impact of stress on how much they want to speculate. However, we are hesitant to take these results

---

[17] The D2 test results consist of the Type-1 and Type-2 errors; we analyze the number of mistakes combined, but the main source were the Type-1 errors.

seriously as they stem from rather low number of observations and this question would need a proper comparison in a separate research paper.

Conclusion and Discussion

In this paper we present the results of a controlled laboratory experiment that shows the existence of a strong gender-specific causal effect of stress on speculative behavior: men under stress speculate more while women speculate less than their non-stressed controls, and the channel seems to be the physical stress response induced by the treatment as measured by the cortisol increase. Furthermore, we have shown that the effects seem to be only very little driven by the treatment-induced change in abilities for k-level thinking and strategic expectations, while task-specific risk-preferences and loss of attention seem to play no role at all. Furthermore, the speculative behavior seems not to be modulated by the trait anxiety, even though the results hold only for the low-anxious subjects.

If the SET score implies behavior in the asset markets (as in Janssen et al., 2015), since we find that men under stress speculate more, the reduction of stress traders undergo might lead to the reduction of the amplitudes in the markets. Our results for men further support the findings of Kocher et al. (2015) who used a standard psychological ego-depletion task (the Stroop test) to test the effect on the size of the bubbles in the SSW markets. The depleted subjects created significantly *larger* bubbles than the non-depleted, while their risk-preferences and of the cognitive resources were unaffected. The authors however did not search for any specific effect of gender. We check for the effects of cognitive depletion by measuring the attention of subjects after the experiment and show that the number of mistakes in the test does not explain the variation in the speculative behavior.

Our finding that the change in the behavior is not due to the change in the risk-preferences contributes to the existing literature contradicting some studies (Buckert, Schwieren, Kudielka, & Fiebach, 2014; Cahlíková & Cingl, 2017) that suspected risk-preferences to be one channel of stress effects on the speculative behavior (Fellner & Maciejovsky, 2007), supporting the opinion that the relationship between stress and risk-preferences is more complicated and may depend on the nature and the timing of the task (Pabst, Brand, & Wolf, 2013; Sokol-Hessner et al., 2016; Vinkers et al., 2013). Furthermore, in the context of financial markets the argument of Coates & Herbert (2008) about hormones irrationally amplifying the business cycle through changes in preferences seems according to our results to work well for men, but not for women.

Our results from the gender-mixed sessions suggest that simply increasing share of women would lead to more stability: women in the gender-mixed sessions did not react to the stress-induction, unlike women in the all-female sessions who decreased their speculative behavior. More women in the markets would thus lead to smaller amplitudes caused by pure speculative behavior. We repeat however, that our results in this dimension rest on rather small number of observations and thus should be taken with a grain of salt.

Recently two orthogonal mental dimensions have been proposed to explain the heterogeneous trading styles in the asset markets: the quantitative and the perspective-taking capabilities (Hefti, Heinke, & Schneider, 2016). We focused on the perspective-taking capability since experiments using the Beauty contest suggest the effect of stress on the strategizing abilities (Leder et al., 2013, 2015). If we assume beliefs about speculation of others reflect those, they seem to play only a minor role in the explanation of the speculative behavior in our experiment. We do not measure the quantitative capabilities and assume that the SET is quantitatively not demanding and does not depend on it (as argued in Janssen et al., 2015).

We also contribute to the literature on the heterogeneous stress effects based on trait anxiety STAI-T that has been shown to matter in how stress affects confidence (Goette et al., 2015). Our analysis of the median split of the sample based on the STAI-T scores present that the subjects high in STAI-T do not show any treatment effect and the main results are driven by the low-anxious subjects, which is what should be observed if stress increases confidence of the low-anxious subjects. We did not measure confidence directly, however. We claim that this provides an increased external validity to our general results since it may be argued that the real-world traders are specially self-selected to be good under stress, which may manifest by a lower trait anxiety score than average population, which is what we found in a related study (Cingl & Zajíček 2016, unpublished results).

References

Abdellaoui, M., Bleichrodt, H., & Kammoun, H. (2013). Do financial professionals behave according to prospect theory? An experimental study. *Theory and Decision*, *74*(3), 411–429. http://doi.org/10.1007/s11238-011-9282-3

Allen, A. P., Kennedy, P. J., Cryan, J. F., Dinan, T. G., & Clarke, G. (2014). Biological and psychological markers of stress in humans: Focus on the Trier Social Stress Test. *Neuroscience and Biobehavioral Reviews*, *38*, 94–124. http://doi.org/10.1016/j.neubiorev.2013.11.005

Avery, B. C., & Zemsky, P. (1998). Multidimensional Uncertainty and Herd Behavior in Financial Markets. *The American Economic Review*, *88*(4), 724–748.

Blanchard, O. J., & Watson, M. W. (1982). Bubbles, rational expectations, and financial markets. *NBER WORKING PAPER SERIES*, *945*, 295–315. http://doi.org/10.3386/w0945

Brickenkamp, R., & Zillmer, E. (1998). *The d2 Test of Attention*. Göttingen: Hogrefe & Huber Publishers.

Buckert, M., Schwieren, C., Kudielka, B. M., & Fiebach, C. J. (2014). Acute stress affects risk taking but not ambiguity aversion. *Frontiers in Neuroscience*, *8*(May), 82. http://doi.org/10.3389/fnins.2014.00082

Buser, T., Dreber, A., & Mollerstrom, J. (2017). The impact of stress on tournament entry. *Experimental Economics*, *20*, 506–530.

Cahlíková, J., & Cingl, L. (2017). Risk preferences under acute stress. *Experimental Economics*, *20*(1), 209–236. http://doi.org/10.1007/s10683-016-9482-3

Cahlikova, J., Cingl, L., & Levely, I. (2017). How Stress Affects Performance and Competitiveness across Gender. *CERGE-EI Working Paper Series*, (589). http://doi.org/10.1007/s13398-014-0173-7.2

Cavatorta, E., & Schröder, D. (2014). Measuring Ambiguity Preferences. *Working Paper*.

Coates, J. M., & Herbert, J. (2008). Endogenous steroids and financial risk taking on a London trading floor. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16), 6167–72. http://doi.org/10.1073/pnas.0704025105

Costa, J. P. T., & McCrae, R. R. (1992). Normal Personality Assessment in Clinical Practice: The NEO Personality Inventory. *Psychological Assessment*, *4*(1), 5–13. Retrieved from http://www.sciencedirect.com/science/article/B6WYX-46VD04F-2/2/430de88263303f8e481e3dac5a7a52b7

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448–474. Retrieved from http://www.ingentaconnect.com/content/aea/jel/2009/00000047/00000002/art00003

Cueva, C., Roberts, R. E., Spencer, T., Rani, N., Tempest, M., Tobler, P. N., … Rustichini, A. (2015). Cortisol and testosterone increase financial risk taking and may destabilize markets. *Scientific Reports*, *5*, 11206. http://doi.org/10.1038/srep11206

Cueva, C., & Rustichini, A. (2015). Is financial instability male-driven? Gender and cognitive skills in experimental asset markets. *Journal of Economic Behavior and Organization*, *119*, 330–344. http://doi.org/10.1016/j.jebo.2015.08.014

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990a). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, *98*(4), 703–738.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990b). Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *The Journal of Finance*, *45*(2), 379–395.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin*, *130*(3), 355–391. Retrieved from http://www.sciencedirect.com/science/article/pii/S0033290904009451

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. *Journal of the European Economic Association*, *9*(3), 522–550. http://doi.org/10.1111/j.1542-4774.2011.01015.x

Eckel, C. C., & Füllbrunn, S. C. (2015). Thar " SHE " blows ? Gender, Competition, and Bubbles in

Experimental Asset Markets. *American Economic Review*, *105*(2), 906–920. http://doi.org/10.1257/aer.20130683

Eckel, C. C., & Füllbrunn, S. C. (2017). Hidden vs. known gender effects in experimental asset markets. *Economics Letters*, *156*, 7–9. http://doi.org/10.1016/j.econlet.2017.04.009

Everly, G. S., & Lating, J. M. (2013). The Anatomy and Physiology of the Human Stress Response. In *A Clinical Guide to the Treatment of the Human Stress Response* (pp. 17–51). http://doi.org/10.1007/978-1-4614-5538-7

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2016). The Preference Survey Module: A Validated Instrument for Measuring Risk , Time , and Social Preferences, (9674).

Fellner, G., & Maciejovsky, B. (2007). Risk attitude and market behavior: Evidence from experimental asset markets. *Journal of Economic Psychology*, *28*(3), 338–350. http://doi.org/10.1016/j.joep.2007.01.006

Filippin, A., & Crosetto, P. (2016). A Reconsideration of Gender Differences in Risk Attitudes. *Management Science*, *62*(11), 3138–3160.

Fischbacher, U. (2007). Z-TREE: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178. Retrieved from http://econpapers.repec.org/RePEc:kap:expeco:v:10:y:2007:i:2:p:171-178

Goette, L., Bendahan, S., Thoresen, J., Hollis, F., & Sandi, C. (2015). Stress pulls us apart : Trait anxiety modulates self-confidence under stress. *Psychoneuroendocrinology*, *54*, 115–123. http://doi.org/10.1016/j.psyneuen.2015.01.019

Greiner, B. (2004). An online recruitment system for economic experiments. In K. Kremer & V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen GWDG Bericht*. Göttingen: Gesellschaft fur Wissenschaftliche Datenverarbeitung.

Hariharan, G., Chapman, K. S., & Domian, D. L. (2000). Risk tolerance and asset allocation for investors nearing retirement. *Financial Services Review*, *9*, 159–170.

Haushofer, J., Cornelisse, S., Seinstra, M., Fehr, E., Joëls, M., & Kalenscher, T. (2013). No effects of psychosocial stress on intertemporal choice. *PloS One*, *8*(11), e78597. http://doi.org/10.1371/journal.pone.0078597

Hefti, A., Heinke, S., & Schneider, F. (2016). Mental Capabilities , Trading Styles, and Asset Market Bubbles: Theory and Experiment. *University of Zurich Working Paper*, (234).

Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, *34*, 163–171. http://doi.org/10.1016/j.psyneuen.2008.10.026

Holt, C. A., Porzio, M., & Yingze Song, M. (2017). Price Bubbles, Gender, and Expectations in Experimental Asset Markets. *European Economic Review*. http://doi.org/10.1016/j.euroecorev.2017.05.005

Hong, H., & Stein, J. C. (1999). A Unified Theory of Underreaction , Momentum Trading , and Overreaction in Asset Markets. *Journal of Finance*, *54*(6), 2143–2184.

Charness, G., & Gneezy, U. (2012). Strong Evidence for Gender Differences in Risk Taking. *Journal of Economic Behavior & Organization*, *83*(1), 50–58. http://doi.org/http://dx.doi.org/10.1016/j.jebo.2011.06.007

Charness, G., Rigotti, L., & Rustichini, A. (2007). Individual Behavior and Group Membership. *American Economic Review*, *97*(4), 1340–1352.

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *The American Economic Review*, *99*(1), 431–457. Retrieved from http://www.jstor.org/stable/10.2307/29730190

Janssen, D.-J., Weitzel, U., & Füllbrunn, S. C. (2015). *Speculative Bubbles - An introduction and application of the Speculation Elicitation Task (SET)*.

Jianakoplos, N. A., & Bernasek, A. (1998). Are Women More Risk Averse. *Economic Inquiry*, *36*(4), 62–630. http://doi.org/10.1111/j.1465-7295.1998.tb01740.x

Kajantie, E., & Phillips, D. I. W. (2006). The effects of sex and hormonal status on the physiological response to acute psychosocial stress. *Psychoneuroendocrinology*, *31*(2), 151–178. http://doi.org/10.1016/j.psyneuen.2005.07.002

Kandasamy, N., Hardy, B., Page, L., Schaffner, M., Graggaber, J., Powlson, A. S., … Coates, J. M. (2014). Cortisol shifts financial risk preferences. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3608–13. http://doi.org/10.1073/pnas.1317908111

Kemeny, M. E. (2003). The Psychobiology of Stress. *Current Directions in Psychological Science*, *12*(4), 124–129. http://doi.org/10.1111/1467-8721.01246

Kirschbaum, C., Pirke, K., & 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. Retrieved from http://www.karger.com/DOI/10.1159/000119004

Kocher, M. G., Lucks, K. E., & Schindler, D. (2015). Unleashing Animal Spirits - Self-Control and Overpricing in Experimental Asset Markets. *CESifo Working Paper*, *49*(0), 1–51.

Leder, J., Häusser, J. A., & Mojzisch, A. (2013). Stress and strategic decision-making in the beauty contest game. *Psychoneuroendocrinology*, *38*(9), 1503–11. http://doi.org/10.1016/j.psyneuen.2012.12.016

Leder, J., Häusser, J. A., & Mojzisch, A. (2015). Exploring the underpinnings of impaired strategic decision-making under stress. *Journal of Economic Psychology*. http://doi.org/10.1016/j.joep.2015.05.006

List, J. A., & Haigh, M. S. (2010). Investment under Uncertainty: Testing the Options Model with Professional Traders. *NBER WORKING PAPER SERIES*, (16038).

Lo, A. W., & Repin, D. V. (2002). The Psychophysiology of Real-Time Financial Risk Processing. *Journal of Cognitive Neuroscience*, *14*(3), 323–339. Retrieved from http://econpapers.repec.org/RePEc:nbr:nberwo:8508

Michailova, J. (2011). *Overconfidence and bubbles in experimental asset markets* (Munich Personal RePEc Archive No. 30579). Munich.

Moinas, S., & Pouget, S. (2013). The Bubble Game: An Experimental Study of Speculation. *Econometrica*, *81*(4), 1507–1539. http://doi.org/10.3982/ECTA9433

Nadler, A., Jiao, P., Johnson, C. J., Alexander, V., & Zak, P. J. (2017). The Bull of Wall Street: Experimental Analysis of Testosterone and Asset Trading. *Management Science*, (forthcoming), 1–53. http://doi.org/10.1287/mnsc.2017.2836

Nicolson, N. (2007). Measurement of cortisol. In L. J. and G. Leucken L. C. (Ed.), *Handbook of physiological research methods in health psychology* (pp. 37–73). SAGE Publications.

Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, *3*(1), 601–630.

Niessen, A., & Ruenzi, S. (2015). Sex Matters : Gender Differences in a Professional Setting. *Working Paper*, (February), 1–2. http://doi.org/10.1007/s10273-011-1262-2

Oberlechner, T., & Nimgade, A. (2005). Work stress and performance among financial traders. *Stress and Health*, *21*(5), 285–293. http://doi.org/10.1002/smi.1063

Pabst, S., Brand, M., & Wolf, O. T. (2013). Stress and decision making: a few minutes make all the difference. *Behavioural Brain Research*, *250*, 39–45. http://doi.org/10.1016/j.bbr.2013.04.046

Palan, S. (2013). a Review of Bubbles and Crashes in Experimental Asset Markets. *Journal of Economic Surveys*, *27*(3), 570–588. http://doi.org/10.1111/joes.12023

Powell, O., & Shestakova, N. (2016). Experimental asset markets: A survey of recent developments. *Journal of Behavioral and Experimental Finance*, *12*, 14–22. http://doi.org/10.1016/j.jbef.2016.08.003

Reschke-hern, A. E., Okerstrom, K. L., Edwards, A. B., & Tranel, D. (2017). Sex and Stress : Men and Women Show Different Cortisol Responses to Psychological Stress Induced by the Trier Social Stress Test and the Iowa Singing Social Stress Test. *Journal of Neuroscience Research*, *114*(November 2016), 106–114. http://doi.org/10.1002/jnr.23851

Riis-vestergaard, M. I., Ast, V. Van, Cornelisse, S., Joëls, M., & Haushofer, J. (2017). The Effect of Hydrocortisone Administration on Intertemporal Choice.

Shleifer, A., & Summers, L. H. (1990). The Noise Trader Approach to Finance. *The Journal of Economic Perspectives*, *4*(2), 19–33.

Schwabe, L., Tegenthoff, M., Höffken, O., & Wolf, O. T. (2010). Concurrent glucocorticoid and noradrenergic activity shifts instrumental behavior from goal-directed to habitual control. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(24), 8190–6. http://doi.org/10.1523/JNEUROSCI.0734-10.2010

Smith, V., Suchanek, G., & Williams, A. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, *56*(5), 1119–1151. Retrieved from http://www.jstor.org/stable/1911361

Sokol-Hessner, P., Raio, C. M., Gottesman, S. P., Lackovic, S. F., & Phelps, E. A. (2016). Acute stress does not affect risky monetary decision-making. *Neurobiology of Stress*, *5*, 19–25. http://doi.org/10.1016/j.ynstr.2016.10.003

Spielberger, C., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-trait anxiety inventory (form Y) (“self-evaluation questionnaire”)*. Consulting Psychologists Press. Retrieved from http://books.google.cz/books?id=OumuGAAACAAJ

Starcke, K., & Brand, M. (2012). Decision making under stress: A selective review. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1228–1248. http://doi.org/10.1016/j.neubiorev.2012.02.003

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *MDBF-Mehrdimensionaler Befindlichkeitsfragebogen*. Göttingen: Hogrefe.

Sutter, M. (2009). Individual behavior and group membership: comment. *American Economic Review*, *99*(5), 2247–2257.

Taylor, S. E. (2006). Tend and Befriend Biobehavioral Bases of Affiliation Under Stress. *Current Directions in Psychological Science*, *15*(6), 273–277.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A. R., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychological Review*, *107*(3), 411.

Vinkers, C. H., Zorn, J. V, Cornelisse, S., Koot, S., Houtepen, L. C., Olivier, B., … Joëls, M. (2013). Time-dependent changes in altruistic punishment following stress. *Psychoneuroendocrinology*, *38*(9), 1467–75. http://doi.org/10.1016/j.psyneuen.2012.12.012

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), 651–60. http://doi.org/10.1177/0956797611431576

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), 514–22. http://doi.org/10.1016/j.psyneuen.2010.08.004

Zhong, S., Shalev, I., Koh, D., Ebstein, R. P., & Chew, S. H. (2017). Competitiveness and Stress. *International Economics Review*, 1–50. Retrieved from http://ssrn.com/abstract=2717459

Tables

Table 1: Elicitation of SET scores. Source: Janssen et al. (2015)

Buy at this (or lower) price	Never	10^0	10^1	10^2	10^3	10^4	10^5	10^6
Probability of being last	0	0	0	23.08	28.57	46.15	57.14	100
SET score	0	1	2	3	4	5	6	7

Note:

Table 2: Summary statistics

	Total	Treatment group		Diff	Rank-sum (p-value)	N
		Control	Treatment			
Sample	(1)	(2)	(3)	(4)	(5)	
		(A) SET score - Same room (*SET_SR*)				
All	3,66	3,65	3,66	-0,01	0,988	208
Male	3,52	3,23	3,8	-0,57	0,035	112
Female	3,82	4,15	3,5	0,65	0,045	96
		(B) SET score - Different room (*SET_DR*)				
All	3,85	3,87	3,84	0,03	0,63	208
Male	3,7	3,54	3,86	-0,32	0,2	112
Female	4,03	4,25	3,81	0,44	0,056	96
		(C) SET score - Computer (*SET_COMP*)				
All	3,63	3,66	3,6	0,06	0,55	208
Male	3,6	3,57	3,63	-0,06	0,888	112
Female	3,67	3,77	3,56	0,21	0,308	96
		(D) Average beliefs - Same room (*beliefs_SR*)				
All	3,49	3,47	3,5	-0,03	0,724	208
Male	3,34	3,19	3,49	-0,3	0,134	112
Female	3,66	3,8	3,52	0,28	0,365	96
		(E) Average beliefs - Different room (*beliefs_DR*)				
All	3,53	3,48	3,59	-0,11	0,533	208
Male	3,41	3,29	3,53	-0,24	0,259	112
Female	3,68	3,7	3,65	0,05	0,689	96

Note: mean SET scores (panels A – C) and mean average beliefs (panels D and E) by treatment and gender. *Treatment* indicates exposure to the stress-inducing procedure of the TSST-G protocol. Differences tested using Wilcoxon rank-sum test.

Table 3: Regression analysis, *SR_SET*

VARIABLES	(1) *SR_SET*	(2) *SR_SET*	(3) *SR_SET*	(4) *SR_SET*	(5) Women	(6) Men
Treatment	0.00962	0.571*	0.572**	0.572**	-0.646*	0.571**
	(0.280)	(0.321)	(0.258)	(0.259)	(0.332)	(0.257)
Female	0.305	0.914***	0.895***	0.894***		
	(0.264)	(0.305)	(0.295)	(0.296)		
Female X Treatment		-1.217**	-1.199***	-1.198***		
		(0.461)	(0.422)	(0.423)		
STAI			-0.00629	-0.0427		
			(0.106)	(0.157)		
STAI X Treatment				0.0673		
				(0.214)		
Constant	3.513***	3.232***	3.232***	3.232***	4.146***	3.232***
	(0.210)	(0.228)	(0.193)	(0.194)	(0.219)	(0.192)
Observations	208	208	207	207	96	112
R-squared	0.010	0.050	0.048	0.049	0.039	0.043

Notes: OLS. Standard errors are clustered at a session level except for columns 5 and 6 where the number clusters was too low and therefore robust standard errors are presented instead. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is *SR_SET* which is the score in the SET when both of the other two traders in the task were from the same room as subjects. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *STAI* is the standardized score of the trait anxiety.

Table 4: Regression analysis: *DR_SET*

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	*DR_SET*	*DR_SET*	*DR_SET*	*DR_SET*	Women	Men
Treatment	-0.0288	0.321	0.320	0.320	-0.438	0.321
	(0.196)	(0.231)	(0.231)	(0.233)	(0.324)	(0.231)
Female	0.335*	0.714**	0.700**	0.699**		
	(0.200)	(0.282)	(0.288)	(0.289)		
Female X Treatment		-0.759*	-0.747*	-0.746*		
		(0.398)	(0.402)	(0.403)		
STAI			0.0542	0.00612		
			(0.0912)	(0.140)		
STAI X Treatment				0.0889		
				(0.185)		
Constant	3.711***	3.536***	3.536***	3.536***	4.250***	3.536***
	(0.154)	(0.167)	(0.166)	(0.167)	(0.228)	(0.167)
Observations	208	208	207	207	96	112
R-squared	0.014	0.032	0.032	0.033	0.019	0.017

Notes: OLS. Standard errors are clustered at a session level except for columns 5 and 6 where the number clusters was too low and therefore robust standard errors are presented instead. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is *DR_SET* which is the score in the SET when both of the other two traders in the task were from another, different room than subjects. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *STAI* is the standardized score of the trait anxiety.

Table 5: Regression analysis: *COMP_SET*

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	*COMP_SET*	*COMP_SET*	*COMP_SET*	*COMP_SET*	Women	Men
Treatment	-0.0673	0.0536	0.0543	0.0544	-0.208	0.0536
	(0.153)	(0.206)	(0.207)	(0.208)	(0.227)	(0.206)
Female	0.0685	0.199	0.172	0.168		
	(0.153)	(0.211)	(0.212)	(0.211)		
Female X Treatment		-0.262	-0.233	-0.233		
		(0.307)	(0.308)	(0.307)		
STAI			-0.0351	-0.152		
			(0.0886)	(0.123)		
STAI X Treatment				0.216		
				(0.178)		
Constant	3.632***	3.571***	3.571***	3.570***	3.771***	3.571***
	(0.123)	(0.137)	(0.137)	(0.138)	(0.161)	(0.137)
Observations	208	208	207	207	96	112
R-squared	0.002	0.005	0.005	0.015	0.009	0.001

Notes: OLS. Standard errors are clustered at a session level except for columns 5 and 6 where the number clusters was too low and therefore robust standard errors are presented instead. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is *COMP_SET* which is the score in SET when the decisions of the other two traders in the task were done by computer. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *STAI* is the standardized score of the trait anxiety.

Table 6: Regression analysis: beliefs SR

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	beliefs_SR	beliefs_SR	beliefs_SR	beliefs_SR	Women	Men
Treatment	0.0316	0.304	0.304	0.304	-0.286	0.304
	(0.165)	(0.211)	(0.211)	(0.211)	(0.258)	(0.211)
Female	0.323*	0.617**	0.682***	0.680***		
	(0.168)	(0.240)	(0.234)	(0.235)		
Female X Treatment		-0.589*	-0.652**	-0.652**		
		(0.333)	(0.329)	(0.330)		
STAI			-0.0437	-0.0824		
			(0.0832)	(0.125)		
STAI X Treatment				0.0717		
				(0.168)		
Constant	3.322***	3.186***	3.186***	3.185***	3.804***	3.186***
	(0.140)	(0.158)	(0.158)	(0.158)	(0.181)	(0.158)
Observations	208	208	207	207	96	112
R-squared	0.018	0.033	0.043	0.044	0.013	0.019

Notes: OLS. Standard errors are clustered at a session level except for columns 5 and 6 where the number clusters was too low and therefore robust standard errors are presented instead. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the average beliefs about SET score of other subjects in the same room, *beliefs_SR*. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *STAI* is the standardized score of the trait anxiety.

Table 7: Regression analysis, beliefs DR

	(1)	(2)	(3)	(4)	(5)	(6)
	beliefs_DR	beliefs_DR	beliefs_DR	beliefs_DR	Women	Men
Treatment	0.106	0.239	0.239	0.239	-0.0495	0.239
	(0.157)	(0.197)	(0.197)	(0.199)	(0.250)	(0.197)
Female	0.267*	0.411*	0.469*	0.466*		
	(0.159)	(0.241)	(0.239)	(0.239)		
Female X Treatment		-0.289	-0.348	-0.347		
		(0.318)	(0.317)	(0.316)		
STAI			0.0203	-0.0747		
			(0.0757)	(0.118)		
STAI X Treatment				0.176		
				(0.153)		
Constant	3.359***	3.292***	3.292***	3.291***	3.703***	3.292***
	(0.128)	(0.143)	(0.143)	(0.145)	(0.194)	(0.143)
Observations	208	208	207	207	96	112
R-squared	0.016	0.020	0.025	0.031	0.000	0.013

Notes: OLS. Standard errors are clustered at a session level except for columns 5 and 6 where the number clusters was too low and therefore robust standard errors are presented instead. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the average beliefs about SET score of other subjects in the different room *beliefs_DR*. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *STAI* is the standardized score of the trait anxiety.

Appendix

Figure 8: SET scores across treatments.

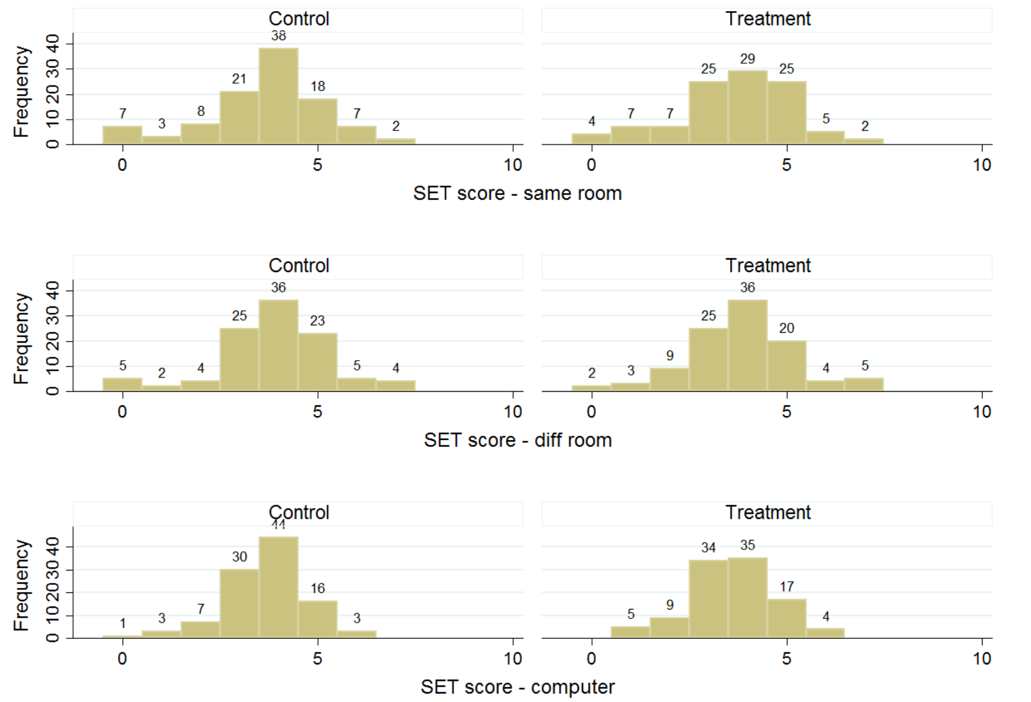
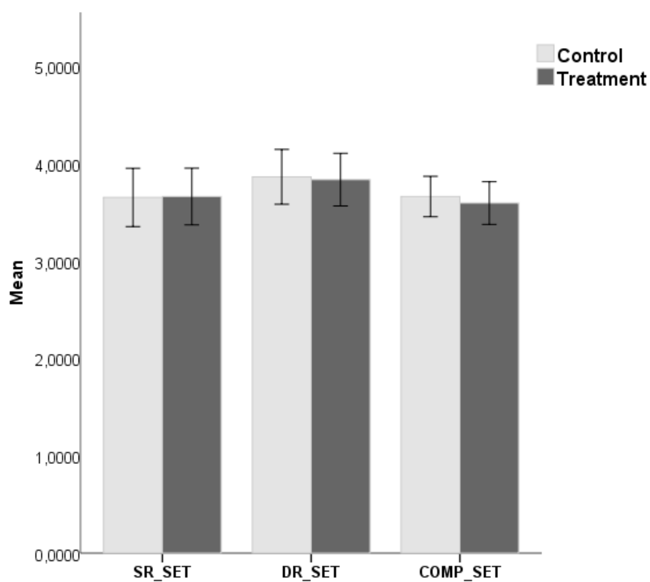


Figure 9: SET scores across treatments.



Note: Darker color indicates the treatment group. Error bars show 95% confidence intervals.

Table 8: Randomization check.

		Pooled			Males			Females		
		Control (n=104)	Treatment (n=104)	Total (n=208)	Control (n=56)	Treatment (n=56)	Total (n=112)	Control (n=48)	Treatment (n=48)	Total (n=96)
Female	Mean	0,46	0,46	0,46						
	SD	0,5	0,5	0,5						
Age	Mean	20,97	21,25	21,11	21,21	21,36	21,29	20,69	21,13	20,91
	SD	1,42	1,28	1,35	1,33	1,3	1,31	1,48	1,25	1,38
Risk 1	Mean	5,36	5,48	5,42	5,57	5,79	5,68	5,11	5,13	5,12
	SD	2,14	2,13	2,13	2,19	2,16	2,17	2,07	2,05	2,05
Ambiguity 1	Mean	3,35	3,4	3,38	3,45	3,21	3,33	3,23	3,63	3,43
	SD	1,15	1,11	1,13	1,19	1,16	1,17	1,11	1,02	1,08
STAI	Mean	39,17	39,63	39,4	37,75	37,91	37,83	40,85	41,65	41,25
	SD	8,45	9,07	8,75	7,46	9,08	8,27	9,3	8,72	8,98
NEO O	Mean	29,05	28,11	28,58	28,8	27,86	28,33	29,33	28,4	28,86
	SD	5,33	6,17	5,77	4,5	5,97	5,28	6,21	6,46	6,32
NEO E	Mean	31,88	32,55	32,21	31,38	32,5	31,94	32,46	32,6	32,53
	SD	6,46	6,48	6,46	6,85	6,94	6,89	5,99	5,96	5,94
NEO C	Mean	31,1	32,69	31,89	31,32	31,09	31,21	30,83	34,56	32,7
	SD	7,58	8,03	7,83	7,21	8,21	7,69	8,06	7,48	7,96
NEO A	Mean	30,06	29,26	29,66	29,86	28,55	29,21	30,29	30,08	30,19
	SD	5,65	6,48	6,07	5,78	6,28	6,05	5,54	6,66	6,09
NEO N	Mean	19,88	18,92	19,4	18,45	16,07	17,26	21,56	22,25	21,91
	SD	8,27	8,65	8,46	7,89	7,9	7,95	8,47	8,38	8,39

Note: *Risk 1*, *Ambiguity 1* and *NEO-O* to *NEO-N* are variables resulting from a non-incentivized internet questionnaire sent to subjects several days prior their participation. *Risk 1* results from the question "*How much are you willing to risk in general?*" on the scale 0 to 10 (Falk, Becker, Dohmen, Huffman, & Sunde, 2016) and Ambiguity 1 results from the following question describing the Ellsberg urn experiment: "*Please imagine the following situation: You can choose between drawing a ball from two different urns, urn A and urn B. Urn A contains 100 balls, some are white and some are black. However, you don't know how many balls are white and how many balls are black. Any combination is possible. There might be from 0 to 100 white balls, with the remaining balls being black. Urn B contains 100 balls as well, but you know that it contains exactly 50 white balls and 50 black balls. Now choose a color, either white or black. Suppose you win £100 if you draw a ball of the color you have selected. If the ball is of the other color, you win nothing. From which urn would you prefer drawing a ball?*
1. I have a strong preference to draw a ball from urn A.
2. I have a slight preference to draw a ball from urn A.
3. I am indifferent between drawing a ball from urn A or from urn B.
4. I have a slight preference to draw a ball from urn B.
5. I have a strong preference to draw a ball from urn B." (Cavatorta & Schröder, 2014)
NEO – O to NEO-N refer to the Big-five personality dimensions Openness to Experience, Extraversion, Conscientiousness, Agreeableness and Neuroticism that were calculated each from 20 questions set in a randomized but fixed order (Costa & McCrae, 1992). F-statistics of the OLS regression where Treatment is the left-hand variable while the rest of the variables are on the right-hand yields 0.99 and none of the variables comes out significant; for men separately the F-statistics is 1.52 and for women separately 1.72.

Table 9: STAI – *SR_SET*

Sample	(1) All	(2) TA-low	(3) TA-high	(4) All	(5) TA-low	(6) TA-high	(7) All	(8) TA-low	(9) TA-high
Treatment	0.00962	-0.132	0.129	0.571*	0.760*	0.367	0.572*	0.776*	0.345
	(0.280)	(0.334)	(0.333)	(0.321)	(0.397)	(0.437)	(0.322)	(0.414)	(0.454)
Female	0.305	0.162	0.427	0.914***	1.120**	0.687	0.895***	1.149**	0.617
	(0.264)	(0.363)	(0.349)	(0.305)	(0.423)	(0.483)	(0.317)	(0.437)	(0.525)
Female X Treatment				-1.217**	-1.870***	-0.530	-1.199**	-1.929***	-0.452
				(0.461)	(0.594)	(0.676)	(0.469)	(0.597)	(0.723)
STAI							-0.00629	0.262	0.0708
							(0.101)	(0.306)	(0.230)
Constant	3.513***	3.719***	3.336***	3.232***	3.240***	3.226***	3.232***	3.436***	3.184***
	(0.210)	(0.256)	(0.271)	(0.228)	(0.252)	(0.311)	(0.229)	(0.285)	(0.323)
Observations	208	103	105	208	103	105	207	103	104
R-squared	0.010	0.005	0.021	0.050	0.109	0.028	0.048	0.116	0.026

Note: OLS regressions. Robust standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is *SR_SET*. Columns 2, 5, 8 feature results run on the subsample of the low-anxious subjects according to the median split, and columns 3, 6, 9 feature results from the high-anxious subsample. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *STAI* is the standardized score of the trait anxiety.

Table 10: STAI – *DR_SET*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample	All	TA-low	TA-high	All	TA-low	TA-high	All	TA-low	TA-high
Treatment	-0.0288	-0.143	0.0980	0.321	0.806**	-0.0980	0.320	0.823**	-0.0939
	(0.154)	(0.283)	(0.221)	(0.211)	(0.338)	(0.266)	(0.212)	(0.359)	(0.275)
Female	0.335**	0.421	0.245	0.714***	1.440***	0.0309	0.700***	1.472***	-0.0166
	(0.150)	(0.320)	(0.238)	(0.182)	(0.370)	(0.275)	(0.186)	(0.389)	(0.290)
Female X Treatment				-0.759***	-1.989***	0.437	-0.747***	-2.054***	0.483
				(0.256)	(0.475)	(0.481)	(0.263)	(0.493)	(0.494)
STAI							0.0542	0.290	-0.0136
							(0.0899)	(0.239)	(0.189)
Constant	3.711***	3.669***	3.747***	3.536***	3.160***	3.839***	3.536***	3.377***	3.847***
	(0.124)	(0.250)	(0.123)	(0.146)	(0.254)	(0.0941)	(0.143)	(0.265)	(0.147)
Observations	208	103	105	208	103	105	207	103	104
R-squared	0.014	0.024	0.009	0.032	0.141	0.016	0.032	0.150	0.017

Note: OLS regressions. Robust standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is *DR_SET*. Columns 2, 5, 8 feature results run on the subsample of the low-anxious subjects according to the median split, and columns 3, 6, 9 feature results from the high-anxious subsample. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *STAI* is the standardized score of the trait anxiety.

Table 11: Physical response to stress: *SR_SET*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:			All		Males		Females	
Estimation:	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Cortisol - pct increase (2-1)	0.00919	0.00756	0.0996	0.425**	0.0996	0.425**	-0.133	-0.544*
	(0.0786)	(0.165)	(0.109)	(0.203)	(0.108)	(0.202)	(0.117)	(0.294)
Cortisol - pct increase (2-1) X Female			-0.232	-0.969***				
			(0.159)	(0.357)				
Female	0.306	0.306	0.414*	0.761***				
	(0.213)	(0.216)	(0.231)	(0.259)				
Constant	3.513***	3.514***	3.465***	3.291***	3.465***	3.291***	3.879***	4.052***
	(0.139)	(0.163)	(0.145)	(0.173)	(0.145)	(0.173)	(0.180)	(0.194)
Observations	208	208	208	208	112	112	96	96
R-squared	0.010	0.010	0.018	-0.059	0.008	-0.078	0.007	-0.064

Note: The dependent variable is *SR_SET*. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *Cortisol – pct increase (2-1)* is the percentage increase in the cortisol levels between the baseline sample collected before the start of the TSST-G (Sample 1) and the sample taken after the first part of the TSST-G procedure (Sample 2). In the IV estimations (columns 6 and 8) the cortisol increase is instrumented by *Treatment* dummy, and in column 4 also with the interaction with Female in the

Table 12: Attention as channel of behavioral change.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All		Males		Females	
Estimation:	OLS	IV	OLS	IV	OLS	IV
Errors D2	0.00303	-0.0494	0.00614*	0.410	-0.0135	-0.263
	(0.00481)	(0.132)	(0.00360)	(1.059)	(0.0228)	(0.192)
Female	0.393*	0.351				
	(0.217)	(0.269)				
Constant	3.417***	3.685***	3.401***	1.343	3.881***	4.956***
	(0.141)	(0.714)	(0.141)	(5.769)	(0.195)	(0.758)
Observations	200	200	104	104	96	96
R-squared	0.017	-0.133	0.004	-17.629	0.003	-0.954

Note: The dependent variable is *SR_SET*. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Treatment is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. Errors D2 is the total number of errors done by subject in the test, type 1 and type 2 errors combined.

Table 13: Mixed sessions.

	(1)	(2)	(3)	(4)	(5)	(6)
		Males			Females	
Treatment	0.571**	0.571**	0.350	-0.646*	-0.646*	-1.062**
	(0.257)	(0.256)	(0.282)	(0.332)	(0.330)	(0.426)
Mixed		-0.375	-0.762*		0.500	-0.125
		(0.310)	(0.451)		(0.328)	(0.478)
Mixed X Treatment			0.775			1.250*
			(0.615)			(0.640)
Constant	3.232***	3.339***	3.450***	4.146***	3.979***	4.187***
	(0.192)	(0.198)	(0.210)	(0.219)	(0.243)	(0.265)
Observations	112	112	112	96	96	96
R-squared	0.043	0.058	0.074	0.039	0.059	0.092

Note: The dependent variable is *SR_SET*. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. *Treatment* is a dummy indicating exposure to the stress-inducing procedure of the TSST-G protocol. *Mixed* indicates if the session was composed of 50 % males and 50 % females, otherwise these were single-sex sessions.