Electronic Supplementary Material (ESM)

**Research Methods**

**Participants**

*Sample size*

The sample size calculation was based on the results of the study with a comparable design in healthy participants which reported an effect size of d=0.77 [1]. A power analysis using this effect size yielded that 16 participants per condition and per group are needed with a power of 0.8 and a two-sided alpha of 0.05 as determined by G\*Power software [2]. Male and female participants were equally represented over the groups and conditions, where possible.

*Recruitment*

Participants were recruited through the advertisements in local newspapers and through an online platform Hersenonderzoek.nl ([www.hersenonderzoek.nl](http://www.hersenonderzoek.nl)).

*Exclusion criteria*

Exclusion criteria for both healthy subjects and patients were: 1) use of insulin or insulin stimulating medications; 2) use of medication that influences glucose metabolism (for example, corticosteroid medication, chemotherapy, beta-blockers); 3) diagnosis of a chronic non-communicable disease (degenerative diseases, malignant neoplasms such as cancer, diabetes type-1, auto-immune diseases); 4) diagnosis of an acute infectious disease (such as meningitis, hepatitis B, bacterial pneumonia); 5) current diagnosis of a mental disorder; 6) chronic and/or acute rhinitis, 7) anatomic deviations of the nose; 8) substance abuse (e.g., drugs or alcohol); 9) pregnancy.

*Randomization*

Randomization was performed by the by the Department of Clinical Pharmacy of the Leiden University Medical Center. The block randomization was used with a size of a block of eight participants per block. Equal numbers of males and females were randomized to each condition.

**Materials**

*Insulin spray*

The unconditioned stimulus was 20 units (0.2 ml) of fast-acting insulin (Insulin NovoRapid; Novo Nordisk), administered with the MAD Nasal Intranasal Mucosal Atomization Device (Teleflex). Six administrations of insulin were done on the first test day in the conditioned group with a break of 15 minutes between the administrations. The spray was administered alternating between the left and then right nostrils. The same dosage of insulin has been successfully used in previous research on insulin conditioning in healthy volunteers [1].

*Placebo spray*

Placebo nasal spray was used in the control group and on day 2 in both groups. The spray was prepared by the Department of Clinical Pharmacy of the Leiden University Medical Center. Because of unavailability of meta-cresol, the preservative that gives the particular smell to the insulin nasal spray, another preservative, chlorobutanol, was used to add a smell to the placebo.

*Conditioned stimulus*

A smell of rosewood oil was used as a CS. The oil was purchased online from www.aromaolie.nl. This aroma oil has previously been used successfully in a study on classical conditioning of oxytocin [3], by our study group, and, mixed with peppermint oils, in previous research on conditioning of insulin responses [4, 5]. This smell has been rated as pleasant but unfamiliar in previous research [3]. Commercially available felt-tip pens were filled with rosewood oil. Each pen was closed with a plastic cap to prevent the odor from escaping. During the smell presentation, participants were asked to hold the pen with the cap removed on a distance of approximately 1 cm in front of both nostrils for one minute immediately before and one minute immediately after the nose spray administration. A new pen was used per each participant.

**Measurements**

***Secondary outcomes***

*Hunger* was measured with a self-rated question “How hungry do you feel at the moment”. Participants were asked to give an answer on a 11-point numeric rating scale (0- “not hungry at all”; 10- “the worst hunger I have ever experienced”). Hunger was measured at the beginning of each session, 5 minutes after each spray administration and 20 minutes after the last spray administration.

*Approach tendencies towards food* were measured with a mobile approach avoidance task. Participants were presented with pictures of objects and food items on a smartphone screen. In congruent blocks of the task, they were asked to pull the food items and push the objects. In incongruent blocks of the task, they were asked to do the opposite—to pull the objects and push away the food items. In total, 80 photos of food and 40 photos of objects were presented in a randomized order. During each response, the phone tracked the gravity- and rotation-corrected acceleration of the movement in the direction perpendicular to the face of the screen (100Hz sampling rate). Based on this acceleration two outcome measures were calculated: reaction times (the time between the stimulus presentation and start of response) and force (peak acceleration) [6]. The pictures for the task were taken from the Food Pics Database [7]. The task was presented to the participants on both day 1 and day 2 after the last blood draw.

*Food consumption* was measured with a bogus taste test adapted from previous studies [8, 9]. At the end of day 1 and 2, participants were offered several snacks: nuts, cucumbers, blueberries, tomatoes, red pepper and carrots. They were allowed to eat as much as they wanted to. Afterwards, the weight of the eaten snacks was measured and the total number of calories eaten was calculated. This task was used to measure the food consumption in previous research [9].

*Memory* was assessed by the Auditory Verbal Learning Test in which 15 words were read to participants 5 times and participants were asked to repeat all the words they could remember after each reading. Fifteen minutes after the first assessment participants were asked to name the words, they still were able to recall. This is a reliable test for measuring learning and memory [10]. Immediate recall scores were calculated by summing the number of all correctly recalled words during the first 5 assessments. Learning scores were calculated by subtracting the number of the words successfully recalled on the first assessment from the number of the words recalled during the fifth assessment. Percent of forgetting scores were calculated by subtracting the number of words recalled on the delayed recall task from the number of words recalled on the fifth assessment. Version A of the task was given to participants after the last spray administration of day 1 and version B of the task after the last spray administration of day 2.

*Body mass index* calculated by dividing the weight in kilograms by the square of the height in meters. Weight and height of participants were measured at the start of the day 1.

**Statistical analysis**

*Mixed effects models*

The multilevel structure of the data was defined by measurement time (level 1) nested in participants (level 2). Parameters were estimated using the full maximum likelihood procedure. In all models, the intercept was allowed to vary randomly across participants. Random slopes did not improve the fit of the models and, therefore, they were removed from the final analysis. The assumption of linearity was checked for each model by plotting the model residuals versus the predictor, and visually inspecting the plots. Homogeneity of variance was checked by Levene’s test. Each model was also checked for the normal distribution of its residuals by looking at QQ plots created with Lattice package. In case of violation of any of the assumptions, the data were transformed. The following variables were transformed due to the violation of the homogeneity of variance and non-normal distribution of the residuals: logarithmic transformation was applied to glucose levels of the day 2, C-peptide levels of the day 1 and day 2, the square root transformation was applied to the insulin levels of day 2, inversion transformation was applied to the reaction time in the approach-avoidance task

**Results**

***Secondary outcomes***

*Approach tendencies towards food*

There was a significant effect of the stimulus type (food vs object) and movement time (pull versus push) interaction on the reaction time, indicating that participants in both groups and conditions had a general approach tendency towards food (B=0.126, SE= 0.52, p= 0.017). There was no effect of condition (B= .01, SE= 0.06, p= 0.915) and any interactions between condition and other variables found (all p’s > 0.54). In the model with force as the outcome, no effect of condition (B= -0.961, SE= 1.63, p= 0.558) on the force was found. No interactions between condition and other variables were found (all p’s > 0.102).

**Bogus taste test**

There was no difference between the conditioned and control group (day 1: F(3,62)=.75, p=.392, ηp2=.013; day 2: F(3,62)=.75, p=.392, ηp2=.013) and between patients and healthy controls (day 1: F(3,62)=1.10, p=.299, ηp2=.019; day 2: F(3,62)=1.10, p=.299, ηp2=.019) in the amount of calories eaten. Neither was the effect of the interaction between group and condition significant on day 1 (F(1,62)= 1.62, p= .208, ηp2= .027), and day 2 (F(1,62)= 2.07, p= .155, ηp2= .033). The scores are presented in ESM Table 2.

**Memory**

There was no difference between conditioned and control groups and patients and healthy controls in their memory scores on day 1 and day 2 (all p’s > .171). The scores are presented in ESM Table 2 and the results of the analyses of each of the memory scores are presented in ESM Table 3.

**ESM Table 2**. Bogus test and memory scores on day 1 and 2 across groups and conditions with standard errors.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Conditioned group | | Control group | |
|  | Patients | Healthy controls | Patients | Healthy controls |
| Bogus test day 1 (kcal) | 74.99 (15.66) | 74.23 (20.56) | 142.26 (44.79) | 71.59 (17.49) |
| Bogus test day 2 (kcal) | 78.77 (19.72) | 89.62 (23.7) | 127.45 (37.58) | 62.47 (13.95) |
| Immediate recall day 1 | 44.60 (2.74) | 43.00 (2.50) | 40.89 (2.89) | 40.42 (2.50) |
| Immediate recall day 2 | 45.20 (2.80) | 46.25 (2.55) | 40.78 (2.95) | 41.58 (2.55) |
| Learning day 1 | 5.30 (0.56) | 5.33 (0.51) | 5.67 (0.59) | 4.67(0.51) |
| Learning day 2 | 5.90 (0.84) | 5.75 (0.76) | 4.67 (0.88) | 5.33 (0.76) |
| Percent forgetting day 1 | 0.21 (0.07) | 0.27 (0.06) | 0.269 (.07) | 0.164 (0.06) |
| Percent forgetting day 2 | 0.24 (0.07) | 0.36 (0.06) | 0.25 (0.07) | 0.22 (0.06) |

**ESM. Table 3**. Results of the factorial ANOVAs comparing groups and conditions on memory scores

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Factor | F | p | ηp2 |
| Immediate recall day 1 | Condition (conditioned versus control) | 0.37 | .544 | .009 |
|  | Group (patients versus healthy controls) | 0.02 | .885 | .001 |
|  | Condition\*Group | 0.003 | .960 | <.001 |
| Immediate recall day 2 | Condition (conditioned versus control) | 1.48 | .231 | .035 |
|  | Group (patients versus healthy controls) | 0.45 | .505 | .011 |
|  | Condition\*Group | 0.22 | .646 | .005 |
| Learning day 1 | Condition (conditioned versus control) | 0.04 | .853 | .001 |
|  | Group (patients versus healthy controls) | 0.62 | .434 | .015 |
|  | Condition\*Group | 1.94 | .171 | .045 |
| Learning day 2 | Condition (conditioned versus control) | 0.45 | .508 | .011 |
|  | Group (patients versus healthy controls) | 0.16 | .691 | .004 |
|  | Condition\*Group | 0.02 | .886 | .001 |
| Percent forgetting day 1 | Condition (conditioned versus control) | 0.19 | .666 | .005 |
|  | Group (patients versus healthy controls) | 0.19 | .663 | .005 |
|  | Condition\*Group | 1.78 | .189 | .043 |
| Percent forgetting day 2 | Condition (conditioned versus control) | 0.88 | .354 | .022 |
|  | Group (patients versus healthy controls) | 0.43 | .515 | .011 |
|  | Condition\*Group | 1.15 | .290 | .029 |

1. Stockhorst U, De Fries D, Steingrueber H-J, Scherbaum WA (2011) Unconditioned and conditioned effects of intranasally administered insulin vs placebo in healthy men: a randomised controlled trial. Diabetologia 54(6):1502–1506

2. Faul F, Erdfelder E, Lang A-G, Buchner A (2007) G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior research methods 39(2):175–191

3. Skvortsova A, Veldhuijzen DS, Pacheco-Lopez G, et al (2019) Placebo effects in the neuroendocrine system: conditioning of the oxytocin responses. Psychosomatic medicine

4. Stockhorst U, Steingrüber H-J, Scherbaum WA (2000) Classically conditioned responses following repeated insulin and glucose administration in humans. Behavioural brain research 110(1–2):143–159

5. Stockhorst U, Gritzmann E, Klopp K, et al (1999) Classical conditioning of insulin effects in healthy humans. Psychosomatic medicine 61(4):424–435

6. Zech HG, Rotteveel M, van Dijk WW, van Dillen LF (2020) A mobile approach-avoidance task. Behavior Research Methods 52(5):2085–2097

7. Blechert J, Meule A, Busch NA, Ohla K (2014) Food-pics: an image database for experimental research on eating and appetite. Frontiers in psychology 5:617

8. van den Akker K, Schyns G, Jansen A (2016) Enhancing inhibitory learning to reduce overeating: Design and rationale of a cue exposure therapy trial in overweight and obese women. Contemporary clinical trials 49:85–91

9. Schakel L, Veldhuijzen DS, Middendorp H van, et al (2018) The effects of a gamified approach avoidance training and verbal suggestions on food outcomes. PloS one 13(7):e0201309

10. Lezak MD, Howieson DB, Loring DW, Fischer JS (2004) Neuropsychological assessment. Oxford University Press, USA