**Supplementary Material**

**Open- and closed-label placebo and nocebo suggestions about a sham transdermal patch**

Stefanie H. Meeuwisa,b, Henriët van Middendorpa,b, Adriana P.M. Lavrijsenc, Dieuwke S. Veldhuijzena,b & Andrea W.M. Eversa,b,d

a Leiden University, Faculty of Social and Behavioural Sciences, Institute of Psychology, Health, Medical and Neuropsychology Unit, Leiden, the Netherlands

b Leiden Institute for Brain and Cognition, Leiden University Medical Center, Leiden, the Netherlands

c Department of Dermatology, Leiden University Medical Center, Leiden, the Netherlands

d Department of Psychiatry, Leiden University Medical Center, Leiden, the Netherlands

**Corresponding author**

Stefanie H. Meeuwis, MSc

Leiden University, Faculty of Social and Behavioural Sciences, Institute of Psychology, Health, Medical and Neuropsychology Unit,

P.O. Box 9555, 2300 RB, Leiden, The Netherlands

0031715274077

s.h.meeuwis@fsw.leidenuniv.nl

**Abbreviations**

AN(C)OVA – Analysis of (co)variance

GLM – General linear model

NRS – Numeric Rating Scales

RMA – Repeared measures ANOVA

PA – Positive affect

PANAS – Positive and Negative Affect Schedule

SEM – Standard error of the mean

STAI-S-s – Spielberger State Trait Anxiety Inventory, State anxiety short scale

SS-10 – Sensitive Scale-10

VS – Verbal suggestions

**1. Supplementary Methods**

**1.1. Materials and Measures**

*1.1.1. Elaboration on the verbal suggestions*

The verbal suggestions that were provided were aimed at eliciting positive or negative expectations about itch induced by histamine iontophoresis. The suggestions were provided in a structured and step-wise manner:

1. ***Information letter.*** Prior to participating, all volunteers received an information letter, in which they were told that the study investigated the effects of a transdermal caffeine patch on cognitive abilities and sensitivity to physical stimuli. The letter contained a brief paragraph with the following information: *“With this study we want to investigate the effects of a new application method for caffeine on cognitive skills. Aa transdermal patch (comparable to a nicotine patch) will be applied, which will release a continuous amount of caffeine. In addition, we want to test the effects of this new application method on the sensitivity to physical stimuli. Caffeine is a common stimulant that can, among other, improve concentration and reduce fatigue. Caffeine is naturally found in various food products (e.g., in chocolate, energy drink, coffee and tea), but research shows that a high intake of these foods can often lead to complaints (e.g., upset stomach, caffeine "dip"). This alternative application method of caffeine may offer a solution”* (translated from Dutch).
2. ***Positive and negative suggestions.*** After baseline measures, participants were first given a short explanation about effects of the patch on cognitive abilities: *“Caffeine is a mild psychoactive substance. It makes you more alert and stimulates your nervous system. Because of this it has a positive effect on your brain. Normally it is contained within drinks such as coffee, but it is also possible to administer caffeine through the skin by using a patch”* (translated from Dutch).

Afterwards, positive or negative verbal suggestions were given (the information was identical but for the valence of certain words, which depended on group allocation): “*We know from research that this way of administering caffeine not only has a positive effect on cognitive skills, but also has an effect on the sensitivity to physical stimuli. Previous research has shown that itching* **[decreases | increases]** *strongly after the application of this patch. This is the case for most people (i.e., about 95% of the examined people). The caffeine makes your skin* **[less | more]** *sensitive to physical stimuli. As such, we expect that you will experience* **[less | more]** *itch compared to the first test”* (translated from Dutch).

1. ***Open-label and closed-label add-on.*** Participants in the closed-label groups received no additional information, while those in the open-label groups were provided with the following: *“I just told you that the patch contains caffeine, and that this influences your skin's sensitivity to physical stimuli. In truth, the patch contains no active substance. What we want to test with this study are precisely the effects of this kind of* **[positive | negative]** *suggestions. We know from previous studies that your experience of itch can be influenced by* **[positive | negative]** *suggestions. We call such an effect a* **[placebo | nocebo]** *effect. For example, the suggestion that the patch* **[decreases | increases]** *itch during the test will actually lead to* **[less | more]** *itch. This is due to various processes in the body, for example itch-***[diminishing | enhancing]** *elements that are produced in the brain. Your brain responds to information about getting a drug in the same way as to actually getting the drug. This also happens when people know that they are getting a* **[placebo | nocebo]***. So even though I told you this, you may experience* **[less | more]** *itch during the test this time.”*

*1.1.2. Elaboration on itch measures: calculated mean itch*

Itch was induced by histamine iontophoresis for 2.5 minutes, during which itch was assessed verbally at: 00:00 min, 00:30 min, 01:00 min, 01:30 min, 2:00 min, and 2:30 min. After 2.5 minutes, mean itch during the test (the main outcome of this study, described in the main text of the manuscript) was immediately assessed and the electrodes were removed. One minute after the test, a short follow-up period started (between 1-4 minutes) with itch measures taken at 1:00, 1:30, 2:00, 2:30, 3:00, 3:30 and 4:00 minutes following iontophoresis. For itch during iontophoresis a mean score was calculated by counting all measures taking during the test and dividing them by 5, the results of which are described in this Supplementary. Likewise, a mean score was calculated for itch during follow-up (1-4 minutes post-iontophoresis).

*1.1.3. Stroop test and Trail Making Test*

As part of the cover story of the patch positively influencing cognition, the Stroop test (1) and Trail Making Test (2,3) were assessed at baseline and following suggestions. Stroop interference scores and percentile scores were calculated adjusted for age, sex, and education level. As a large inter-individual variability in the execution of the Trail Making Test was noted (e.g., on noticing and dealing with mistakes during the test – some participants did not correct mistakes whereas others did, thus causing differences on the time spent taking the test and the associated outcome measure), these data were not analysed.

*1.1.4. Expectations for the cognitive tasks*

Prior to the cognitive tasks (both baseline and post-verbal suggestions) participants were asked how well they expected to perform during the tasks by rating the following items on a Numeric Rating Scale (NRS) ranging from 0 (“not good at all”) to 10 (“very good”): focus, attention, performance, and speed during the tasks. In addition, expected patch efficacy for focus, attention, and speed was rated following the verbal suggestions using the same NRS.

*1.1.5. Wellbeing: State Anxiety and General Wellbeing Scales*

The State Trait Anxiety Inventory – State Anxiety short scales (STAI-S-s; (4)), and seven Numeric Rating Scales (NRS) measuring general wellbeing (relaxed, anxious, serene, agreeable, tense, worried, stressed, see also (5)) on a scale from 0 (‘not at all’) to 10 (‘very much so’) were assessed at four moments during the laboratory session: I. pre-baseline iontophoresis (baseline 1), II. post-baseline iontophoresis (baseline 2), III. post-verbal suggestions (post-VS 1), and IV. at the end of the session (post-VS 2). Of the NRS items, negative items were recoded and a total score was calculated by summing all items (with higher scores reflecting higher general wellbeing). Cronbach’s alpha for state anxiety ranged from .80 to .83 in the current study. For general wellbeing, Cronbach’s alpha ranged .86 to .87.

**1.2. Statistical Analysis**

All analyses were conducted in SPSS 25.0 for Windows (IBM SPSS Inc., Chicago, IL, USA). As was a priori determined, open-label and closed-label groups were first combined to detect differences between the effects of positive verbal suggestions and negative verbal suggestions with optimal power, and second, repeated for the separate open-label context and the separate closed-label context. Normal distribution of the variables, baseline differences, and assumptions were checked prior to data analysis. General linear model (GLM) analyses of covariance (ANCOVAs) were used to assess group differences in calculated mean itch during iontophoresis, calculated mean itch during follow-up, and expectations regarding the cognitive tasks. For the expected efficacy of the patch, GLM analysis of variance (ANOVA) was used. Effects of suggestions on Stroop scores were analysed using GLM ANCOVA, and effects of group on wellbeing scales were assessed by mixed between-within-subject RMA. The critical alpha used was α=.05 for the combined group analyses. For the separate open-label and closed-label context analyses, a Bonferroni correction for multiple comparisons was applied (α/2=.025).

**E2. Supplementary Results**

**2.1. Effects of suggestions on itch**

*2.1.1. Calculated mean itch during iontophoresis*

No differences were found for calculated mean itch at the baseline assessment (pre-suggestions iontophoresis) for either combined or separate open- and closed-label contexts (all *p*≥.75). Following verbal suggestions, calculated mean itch during iontophoresis was significantly lower in the combined positive VS groups (*M* = 2.94±1.48) compared to the combined negative VS groups (*M* = 3.62±1.85; see also **Supplementary Figure S2**); F(1,108)=11.67, *p*=.001, Cohen’s d=0.66. For the open-label context, significantly lower calculated mean itch during iontophoresis was found for the positive VS groups (*M* = 2.95±1.62) relative to the negative VS groups (*M* = 3.59±1.76); F(1,52)=6.58, *p*=.013, Cohen’s d=0.71. For the closed-label context marginally significant differences between the positive VS group (*M* = 2.94±1.36) and negative VS group (*M=*3.66±1.97) were found (see also **Supplementary Figure S2**); F(1,53)=5.10, *p*=.028, Cohen’s d=0.62[[1]](#footnote-1).

*2.1.2. Calculated mean itch during post-iontophoresis follow-up*

No differences were found for calculated mean itch during follow-up at the baseline assessment (pre-suggestions iontophoresis) for either combined or separate open- and closed-label contexts (all *p*>.40). Following verbal suggestions, calculated mean itch during follow-up was significantly lower in the combined positive VS groups (*M* = 1.79±1.48) compared to the negative VS groups (*M* = 2.52±2.00; see also **Supplementary Figure S3**); F(1,106)=5.89, *p*=.017, Cohen’s d=0.47. A marginally significant1 difference between the positive VS group (*M* = 1.72±1.28) and the negative VS group (*M* = 2.58±2.05) was found for the closed-label context only; F(1,52)=4.54, *p*=.038, Cohen’s d=0.59. For the open-label context, no significant difference between groups was found (see also **Supplementary Figure S3)**; F(1,51)=1.43, *p=*.24, Cohen’s d=0.33.

**2.2. Effects of suggestions for cognitive tasks and wellbeing**

*2.2.1. Expectations and expected patch efficacy for the cognitive tasks*

At baseline, expected attention and performance during the tasks were significantly higher in the combined open- and closed-label negative VS groups (attention: *M*=7.00±1.39; performance: *M*=6.80±1.07) compared to the combined positive VS groups (attention: *M*=6.33±1.56; performance: *M*=6.26±1.44); Fatt..(1,109)=5.55, *p*=.020, Fperf.(1,109)=5.09, *p*=.026; see **Supplementary Table E1**. Marginal differences were found for expected focus (*p*=.052) and expected speed (*p=*.073), respectively. When open-label and closed-label contexts were separated, no baseline differences could be found (all *p*≥.061; see also **Supplementary Table E2**).

When open- and closed-label groups were combined, no group differences were found for any of the post-suggestions expectation measures (all *p*≥.055), with the exception of expected patch influence on speed; F(1,109)=4.11, *p*=.045, Cohen’s d=0.39 (see **Supplementary Table E1**). Participants in the combined negative VS groups expected the patch to be more effective for speed during the cognitive tasks (*M*=3.70±2.38), compared to participants in the combined positive VS groups (*M*=2.86±1.94). No group differences were found in the separate open-label or closed-label context (all *p*≥.091; see **Supplementary Table E2**). Within-group baseline-to-post-VS-change indicated no significant changes in expectations regarding the cognitive tasks following suggestions after applying the Bonferroni correction for multiple comparisons (all *p*≥.042), with the exception of expected speed in the combined positive VS groups. Within these groups, a significant decrease in expected speed was noted (*M*change=-0.48, t(54)=2.54, *p*=.014 (see **Supplementary Table E3**). When open-label and closed-label contexts were separated, no significant changes from baseline to post-VS were noted (all *p*≥.13; see **Supplementary Table E4**).

*2.2.2. Effects on Stroop Test*

The combined positive and the combined negative VS groups did not differ on Stroop interference scores or percentile scores at baseline, or following verbal suggestions (all *p*≥.10). Similar findings were demonstrated when analyses were repeated for the separate open-label and closed-label contexts (all *p*≥.044[[2]](#footnote-2)). Within-group baseline-to-post-VS-change indicated Stroop interference and percentile scores improved in both the combined positive and the combined negative groups (all *p*<.001). When groups were separated, similar reductions in Stroop interference and percentile scores were found for the closed-label negative VS and the open-label positive VS groups (all *p*≤.020). In the open-label negative VS group, no significant change in interference score was found after applying the Bonferroni correction for multiple comparisons (*p*=.039), however, the percentile score did reduce significantly (*p*=.011). Stroop interference and percentile scores did not change in the closed-label positive VS group (both *p*≥.037).

2.2.3. *Subjective wellbeing: State anxiety and General Wellbeing Scales*

No combined-group x time interactions, or main effects of the combined groups, were found for STAI total score or general wellbeing scales (*p*≥.22). Both state anxiety and general wellbeing changed significantly over time (*p*≤.012, see **Supplementary Figure S4A** (state anxiety) and **Supplementary Figure S5A** (general wellbeing)). For state anxiety, the baseline measurement was significantly lower than the second measurement (baseline post-iontophoresis; *p*=.006). No significant differences over time for the other measurements were found (all *p*≥.080). General wellbeing at the final measurement point was significantly higher compared to the second and third measurements (both *p*≤.015). When analyses were conducted for the separate open-label and closed-label contexts, no effects were found for state anxiety (see **Supplementary Figure S4B**). For general wellbeing, the final measurement was significantly higher compared to the baseline post-VS measurement only for the open-label context (see **Supplementary Figure S5B**).

**Concluding note on the effects of open-label and closed-label suggestions on expectations regarding cognition, outcomes of the cognitive tasks and wellbeing**

Overall, verbal suggestions did not induce differences in expectations regarding the cognitive tasks and patch efficacy for such tasks, with the exception of expected speed during the tasks. The participants in the negative suggestions groups expected the patch to be more effective for speed during the cognitive tasks compared to the positive VS groups. It should be noted though that on general, both groups scored low on expected efficacy for speed (mean of 3.70 and 2.86, respectively, on a 0-10 scale). When groups were separated for open-label and closed-label contexts, no differences in expectations were found. Notably, expectations for the cognitive tasks generally did not reduce after suggestions, indicating that the positive suggestions about the tasks may not have impacted participants’ expectations. In the positive VS groups, participants expected to be slower during the cognitive tasks (as reflected by the significant within-subject change from baseline scores). It should be noted though that expectations were assessed at baseline, that is, before any test was conducted. It may be possible that participants found the Stroop test, Trail Making Test, and filler tasks (i.e. Sudoku’s and other puzzles) more challenging than previously anticipated, which may have impacted the scores given post-suggestions. Alternatively, the manner of assessing expectations did not account for (mental) fatigue, and it may be possible that this influenced participants’ assessment of their expected performance. In general, some improvement on Stroop test scores was demonstrated for all groups, which could be due to training effects. Regarding wellbeing, significant changes over time could be demonstrated, but positive and negative verbal suggestions did not significantly impact either state anxiety or general wellbeing. This is in line with the findings for positive affect, which are described in the paper of the current study.

**Supplementary References**

1. Hammes J, *De Stroop kleur-woord test. Handleiding, tweede gewijzigde druk [The Stroop color-word test. Manual, second revised edition]*. 1978, Lisse: Swets & Zeitlinger.

2. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. Nat Protoc 2006;1:2277.

3. Reitan RM. The relation of the trail making test to organic brain damage. J Consult Psychol 1955;19:393.

4. Spielberger CD, *Manual for the State-Trait Anxiety Inventory*. Revised edition ed. 1983, Palo Alto, CA: Consulting Psychologists Press.

5. Meeuwis SH, van Middendorp H, Pacheco-Lopez G, Ninaber MK, Lavrijsen APM, van der Wee N, Veldhuijzen DS & Evers AWM. Antipruritic placebo effects by conditioning H1-antihistamines. Psychosom Med 2019;81;841-50.

6. Misery L, Jean-Decoster C, Mery S, Georgescu V, Sibaud V. A new ten-item questionnaire for assessing sensitive skin: the sensitive scale-10. Acta Derm Venereol 2014;94:635-9.

**E4. Supplementary Tables and Figures**

**Supplementary Table S1.** Means, standard deviations, and analysis of (co)variance (AN(C)OVA) outcomes for expectations and outcomes of the cognitive tests in the combined open- and closed-label positive and negative verbal suggestions (VS) groups.

|  |  |
| --- | --- |
|  |  |
|  | **Combined open- and closed-label contexts** |
|  |  | **AN(C)OVA** |
|  | **Positive VS (*n*=55)** | **Negative VS (*n*=56)** | ***p*-value** | **Cohen’s d** |
| *Baseline expectation outcomes for cognitive tasks* |  |  |  |  |
| Expected focus | 6.47 ± 1.48 | 6.97 ± 1.17 | .052 | 0.38 |
| Expected attention | 6.33 ± 1.56 | 7.00 ± 1.39 | .020 | 0.45 |
| Expected performance | 6.26 ± 1.44 | 6.80 ± 1.07 | .026 | 0.43 |
| Expected speed | 5.73 ± 1.71 | 6.25 ± 1.29 | .073 | 0.34 |
|  |  |  |  |  |
| *Baseline Stroop test* |  |  |  |  |
| Interference score | 56.67 ± 7.80 | 55.38 ± 8.45 | .40 | 0.16 |
| Percentile score | 69.91 ± 20.87 | 65.41 ± 22.26 | .28 | 0.21 |
|  |  |  |  |  |
| *Post-VS expectation outcomes for cognitive tasks* |  |  |  |  |
| Expected focus | 6.33 ± 1.40 | 6.99 ± 1.41 | .10 | 0.27 |
| Expected attention  | 6.42 ± 1.60 | 7.11 ± 1.33 | .21 | 0.20 |
| Expected performance | 5.93 ± 1.64 | 6.57 ± 1.37 | .29 | 0.16 |
| Expected speed | 5.24 ± 2.00 | 6.13 ± 1.43 | .055 | 0.28 |
| Expected patch influence on focus | 3.75 ± 2.19 | 3.81 ± 2.29 | .88 | 0.03 |
| Expected patch influence on attention | 3.70 ± 2.28 | 3.78 ± 2.10 | .84 | 0.04 |
| Expected patch influence on speed | 2.86 ± 1.94 | 3.70 ± 2.38 | .045 | 0.39 |
|  |  |  |  |  |
| *Post-VS Stroop test* |  |  |  |  |
| Interference score | 61.93 ± 9.54 | 58.86 ± 7.86 | .10 | 0.27 |
| Percentile score | 80.87 ± 22.59 | 75.82 ± 19.00 | .43 | 0.13 |
|  |  |  |  |  |

Note. Cohen’s d for ANCOVA was calculated using the covariate-adjusted means (not depicted in the table).

**Supplementary Table S2.** Means, standard deviations, and analysis of (co)variance (AN(C)OVA) outcomes for expectations and outcomes of the cognitive tests in the separate open-label and closed-label positive and negative verbal suggestions (VS) groups.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  | **Open-label context** |  | **Closed-label context** |
|  |  | **AN(C)OVA** |  |  | **AN(C)OVA** |
|  | **Positive VS (*n*=27)** | **Negative VS (*n*=28)** | ***p*-value** | **Cohen’s d** |  | **Positive VS (*n*=28)** | **Negative VS (*n*=28)** | ***p*-value** | **Cohen’s d** |
| *Baseline expectation outcomes for cognitive tasks* |  |  |  |  |  |  |  |  |  |
| Expected focus | 6.26 ± 1.60 | 6.82 ± 1.33 | .17 | 0.38 |  | 6.68 ± 1.34 | 7.12 ± 0.97 | .16 | 0.38 |
| Expected attention | 6.14 ± 1.59 | 6.75 ± 1.51 | .15 | 0.39 |  | 6.53 ± 1.54 | 7.24 ± 1.25 | .061 | 0.51 |
| Expected performance | 6.16 ± 1.45 | 6.76 ± 1.06 | .084 | 0.47 |  | 6.35 ± 1.45 | 6.83 ± 1.09 | .16 | 0.37 |
| Expected speed | 5.71 ± 1.58 | 6.03 ± 1.41 | .44 | 0.21 |  | 5.74 ± 1.86 | 6.46 ± 1.14 | .084 | 0.47 |
|  |  |  |  |  |  |  |  |  |  |
| *Baseline Stroop test* |  |  |  |  |  |  |  |  |  |
| Interference score | 57.67 ± 7.32 | 55.82 ± 7.67 | .37 | 0.25 |  | 55.71 ± 8.25 | 54.93 ± 9.29 | .74 | 0.09 |
| Percentile score | 73.48 ± 18.96 | 67.00 ± 19.97 | .22 | 0.33 |  | 66.46 ± 22.36 | 63.82 ± 24.60 | .68 | 0.11 |
|  |  |  |  |  |  |  |  |  |  |
| *Post-VS expectation outcomes for cognitive tasks* |  |  |  |  |  |  |  |  |  |
| Expected focus | 6.09 ± 1.53 | 6.87 ± 1.45 | .16 | 0.35 |  | 6.57 ± 1.25 | 7.11 ± 1.37 | .41 | 0.18 |
| Expected attention | 6.34 ± 1.80 | 7.05 ± 1.30 | .34 | 0.21 |  | 6.50 ± 1.40 | 7.16 ± 1.39 | .45 | 0.17 |
| Expected performance | 5.91 ± 1.79 | 6.68 ± 1.20 | .35 | 0.21 |  | 5.94 ± 1.52 | 6.46 ± 1.55 | .61 | 0.11 |
| Expected speed | 5.29 ± 2.22 | 6.12 ± 1.25 | .13 | 0.32 |  | 5.20 ± 1.80 | 6.14 ± 1.62 | .27 | 0.23 |
| Expected patch influence on focus | 3.00 ± 2.22 | 3.19 ± 2.48 | .77 | 0.08 |  | 4.47 ± 1.93 | 4.44 ± 1.92 | .95 | 0.02 |
| Expected patch influence on attention | 2.86 ± 2.11 | 3.20 ± 2.21 | .56 | 0.16 |  | 4.51 ± 2.17 | 4.36 ± 1.84 | .79 | 0.08 |
| Expected patch influence on speed | 2.18 ± 1.79 | 2.98 ± 2.53 | .18 | 0.37 |  | 3.52 ± 1.89 | 4.42 ± 2.02 | .091 | 0.46 |
|  |  |  |  |  |  |  |  |  |  |
| *Post-VS Stroop test* |  |  |  |  |  |  |  |  |  |
| Interference score | 64.52 ± 9.24 | 59.18 ± 8.31 | .044 | 0.48 |  | 59.43 ± 9.31 | 58.54 ± 7.53 | .80 | 0.06 |
| Percentile score | 86.07 ± 20.17 | 76.68 ± 18.86 | .21 | 0.28 |  | 75.86 ± 24.00 | 74.96 ± 19.46 | .94 | 0.02 |
|  |  |  |  |  |  |  |  |  |  |

Note. The critical alpha used following Bonferroni correction for multiple comparisons was α = (.05/2) = .025. Cohen’s d for ANCOVA was calculated using the covariate-adjusted means (not depicted in the table).

**Supplementary Table S3.** Within-group baseline-to-post-verbal suggestions (VS) changes for expectations and outcomes of the cognitive tests in the combined open- and closed-label positive and negative verbal suggestions (VS) groups.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  | **Combined open- and closed- label positive VS groups (*n*=55)** |  | **Combined open- and closed-label negative VS groups (*n*=56)** |
|  | *n* | Mean change | *t* | *p* |  | *n* | Mean change | *t* | *p* |
|  |  |  |  |  |  |  |  |  |  |
| *Expectation outcomes for cognitive tasks* |  |  |  |  |  |  |  |  |  |
| Expected focus | 55 | -0.14 | 0.78 | .44 |  | 56 | 0.03 | -0.15 | .89 |
| Expected attention | 55 | 0.09 | -0.45 | .65 |  | 56 | 0.11 | -0.70 | .49 |
| Expected performance | 55 | -0.33 | 2.09 | .042 |  | 56 | -0.23 | 1.32 | .19 |
| Expected speed | 55 | -0.48 | 2.54 | .014 |  | 56 | -0.12 | 0.68 | .50 |
|  |  |  |  |  |  |  |  |  |  |
| *Stroop test* |  |  |  |  |  |  |  |  |  |
| Interference score | 55 | 5.26 | -4.70 | < .001 |  | 56 | 3.48 | -3.30 | .002 |
| Percentile score | 55 | 10.96 | 3.95 | < .001 |  | 56 | 10.41 | -4.01 | < .001 |
|  |  |  |  |  |  |  |  |  |  |

Note. Mean change was calculated as post-verbal suggestions score – baseline score, with negative values indicating a decrease in scores from baseline, and positive scores indicating an increase in scores from baseline.

**Supplementary Table S4.** Within-group baseline-to-post-verbal suggestions (VS) changes for expectations and outcomes of the cognitive tests in the separate open-label and closed-label positive and negative verbal suggestions (VS) groups.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Open-label context** |  | **Closed-label context** |
|  | **Positive VS group** |  | **Negative VS group** |  | **Positive VS group** |  | **Negative VS group** |
|  | *n* | Mean change | *t* | *p* |  | *n* | Mean change | *t* | *p* |  | *n* | Mean change | *t* | *p* |  | *n* | Mean change | *t* | *p* |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Expectation outcomes for cognitive tasks* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Expected focus | 27 | -0.17 | 0.57 | .58 |  | 28 | 0.05 | -0.20 | .85 |  | 28 | -0.11 | 0.54 | .59 |  | 28 | -0.003 | 0.02 | .99 |
| Expected attention | 27 | 0.21 | -0.73 | .47 |  | 28 | 0.30 | -1.20 | .24 |  | 28 | -0.03 | 0.09 | .93 |  | 28 | -0.08 | 0.42 | .68 |
| Expected performance | 27 | -0.25 | 1.14 | .26 |  | 28 | -0.09 | 0.35 | .73 |  | 28 | -0.41 | 1.76 | .089 |  | 28 | -0.37 | 1.55 | .13 |
| Expected speed | 27 | -0.43 | 1.62 | .12 |  | 28 | 0.09 | -0.31 | .76 |  | 28 | -0.54 | 1.93 | .064 |  | 28 | -0.33 | 1.49 | .15 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Stroop test* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Interference score | 27 | 6.85 | -4.82 | < .001 |  | 28 | 3.36 | -2.17 | .039 |  | 28 | 3.71 | -2.20 | .037 |  | 28 | 3.61 | -2.47 | .020 |
| Percentile score | 27 | 12.59 | -4.02 | < .001 |  | 28 | 9.68 | -2.75 | .011 |  | 28 | 9.39 | -2.05 | .050 |  | 28 | 11.14 | -2.88 | .008 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Note. Mean change was calculated as post-verbal suggestions score – baseline score, with negative values indicating a decrease from baseline, and positive scores indicating an increase from baseline. The critical alpha used following Bonferroni correction for multiple comparisons was α = (.05/2)/2 = .0125.

**Supplementary Figure S1.** Means + SEMs of the Positive and Negative Affect Schedule (PANAS) subscale ‘positive affect’ and mixed between-within repeated measures ANOVA (RMA) outcomes for (A) the combined open- and closed label positive VS groups and the combined negative VS groups, and (B) the separate groups.

**Supplementary Figure S2.** Numeric Rating Scale (NRS) calculated mean score for itch experienced during histamine iontophoresis for the baseline and post-verbal suggestions (VS) measurements, with the standard error of the mean (SEM) for [A] the combined open- and closed-label positive and negative VS groups, and [B] the separate open-label and closed-label positive and negative VS groups.



****

Critical alpha levels for the **combined group** analysis: \*\*\* p<.001, \*\* p<.01, \* p<.05, † p<.10, n.s. non-significant. Bonferroni-corrected critical alpha levels for the **separate group** analyses: \*\*\* p<.001, \*\* p<.005, \* p<.025, † p<.05, n.s. non-significant.

**Supplementary Figure S3.** Numeric Rating Scale (NRS) calculated mean score for itch experienced during the 1-4 minutes follow-up to iontophoresis for the baseline and post-verbal suggestions (VS) measurements, with the standard error of the mean (SEM) for [A] the combined open- and closed-label positive and negative VS groups, and [B] the separate open-label and closed-label positive and negative VS groups.





Critical alpha levels for the **combined group** analysis: \*\*\* p<.001, \*\* p<.01, \* p<.05, † p<.10, n.s. non-significant. Bonferroni-corrected critical alpha levels for the **separate group** analyses: \*\*\* p<.001, \*\* p<.005, \* p<.025, † p<.05, n.s. non-significant.

**Supplementary Figure S4.** Means + SEMs of the State and Trait Anxiety Inventory (STAI) subscale ‘state anxiety’ and mixed between-within repeated measures ANOVA (RMA) outcomes for (A) the combined open- and closed label positive VS groups and the combined negative VS groups, and (B) the separate groups.

**Supplementary Figure S5.** Means + SEMs of the Numeric Rating Scales (NRS) total score for ‘general wellbeing’ and mixed between-within repeated measures ANOVA (RMA) outcomes for (A) the combined open- and closed label positive VS groups and the combined negative VS groups, and (B) the separate groups.

1. Given a Bonferroni correction of the critical alpha level: α=.05/2 = .025 this effect is non-significant. [↑](#footnote-ref-1)
2. Given a Bonferroni correction of the critical alpha level: α=.05/2 = .025 these effects are non-significant. [↑](#footnote-ref-2)