

Deviations from the pre-registration

of the manuscript entitled “Be Yourself and Behave Appropriately: Exploring Associations Between Incongruent Personality States and Positive Affect, Tiredness, and Cognitive performance”

Overview

| Components | Were there deviations? (no, minor, major) | If yes – describe the deviation(s) | Rationale for deviation | How might the results be different if you had not deviated |
|-------------------|---|------------------------------------|---------------------------|--|
| Study Information | Minor (Hypotheses) | Described in detail below | Described in detail below | We do not expect that the results would have been different. |
| Sampling Plan | Minor (Stopping rule) | Described in detail below | Described in detail below | We do not expect that the results would have been different. |
| Variables | No | | | |
| Design Plan | No | | | |
| Analysis Plan | Minor (Statistical models) Minor (Transformations) Minor (Inference criteria) | Described in detail below | Described in detail below | We do not expect that the results would have been different. |
| Scripts | No | | | |
| Other | No | | | |

Hypotheses

| Components | Preregistered Analysis | Reported Analysis | Rationale for deviation | How might the results be different if you had not deviated |
|--|---|--|--|--|
| Inference criteria <i>(also listed under analysis plan, as it normally belongs there)</i> | Thus, when talking about congruence between personality trait and personality state or between personality state and situation characteristic, we will each test these 4 constellations. Therefore, we will use a Bonferroni-corrected significance level of $\alpha = .0125$ for these analyses. | We used a Bonferroni-corrected significance level of $\alpha = .00625$ for these analyses. | As H4c (associations between state-situation congruence and Stroop performance) was tested in 8 separate analyses (not 4), we tightened the significance level to .00625 | We do not expect that the results would have been much different. If at all, our conclusions were now a bit more conservative. |

Sampling Plan

| Components | Preregistered Analysis | Reported Analysis | Rationale for deviation | How might the results be different if you had not deviated |
|-------------------|--|--|---|---|
| Stopping rule | We aim to collect 200 data points with valid data (i.e., questionnaire has been completed and instructed response items have been answered correctly). Based on prior experience with exclusions due to careless responding, we will collect 10% more data, that is, data from 220 participants. | We implemented a filter in the data collection tool that automatically estimated the number of participants with valid data and stopped when this number had reached 200 participants. | It was more precise to monitor the number of valid data points than to estimate the number of potential exclusions. | We do not expect that the results would have been different because the result, the sample size, is the same – we simply took a different approach to achieve it. |

Analysis Plan

| Components | Preregistered Analysis | Reported Analysis | Rationale for deviation | How might the results be different if you had not deviated |
|--------------------|---|---|---|--|
| Inference criteria | Thus, when talking about congruence between personality trait and personality state or between personality state and situation characteristic, we will each test these 4 constellations. Therefore, we will use a Bonferroni-corrected significance level of $\alpha = .0125$ for these analyses. | We used a Bonferroni-corrected significance level of $\alpha = .00625$ for these analyses. | As H4c (associations between state-situation congruence and Stroop performance) was tested in 8 separate analyses (not 4), we tightened the significance level to .00625 | We do not expect that the results would have been much different. If at all, our conclusions were now a bit more conservative. |
| Transformations | We will compute scale scores by reverse-coding items when applicable and then averaging over all items that belong to a scale. This applies to the personality traits, personality states, situation perception, and mood. | We additionally transformed personality traits from a 5-point scale to a 7-point scale for the response surface analyses. We used the following formula to do so: $\text{New} = 6 * (\text{old} - 1) / 4 + 1$ | For response surface analyses, the two constructs must be measured on comparable scales. Thus, equal numerical values should also have the same contextual meaning. Therefore, we had to transform the constructs to lie on the same scale. | We do not expect that the results would have been much different. |

| Components | Preregistered Analysis | Reported Analysis | Rationale for deviation | How might the results be different if you had not deviated |
|---|--|---|---|---|
| <p>Statistical models: Manipulation check</p> | <p>The behavior manipulation will be considered successful (H1a + b) if the behavior condition variable significantly predicts behavior in the game (i.e., number of share decision and/or number of honest trials) and personality states (i.e., state Agreeableness and/or state Honesty-Humility) but not situation perception (i.e., Deception and Adversity of the situation).</p> <p>The situation manipulation will be considered successful (H1c) if the situation condition variable significantly predicts situation perception (i.e., Deception and/or Adversity of the situation) but not personality states (i.e., state Agreeableness and state Honesty-Humility).</p> | <p>Both manipulations were only considered successful if they the intended effects as preregistered AND also did not have unintended effect on the remaining personality states or situation characteristics.</p> | <p>We did not preregister a decision regarding the success of the manipulations in case they also affected the unintended variables.</p> <p>However, we believe that a manipulation that has diffuse effect on both the intended variables and other, unintended variables cannot be considered truly successful. We therefore concluded that the manipulations were unsuccessful in our study.</p> | <p>We do not expect that the results would have been much different.</p> <p>We would have performed more analyses if we had not deviated, that is, additional variance analyses would have been conducted. However, we believe that the conclusions from these models would have been the same.</p> |

| Components | Preregistered Analysis | Reported Analysis | Rationale for deviation | How might the results be different if you had not deviated |
|---|--|--|--|--|
| Statistical models: Specification curve analyses | In sum, we will analyze a total of 2896 RT specifications and 724 error rate specifications. | We examined 2224 reaction time specifications and 556 error rate specifications. | We miscalculated the number of specifications in the preregistration because not all decisions could be meaningfully applied to all indices. | We do not expect that the results would have been different. The specifications curve analyses produced very clear results and a few more specifications would not have changed the conclusions. |