Personality traits describe how people typically think, feel, and behave. Personality states describe how people think, feel, and behave in a given moment. In their daily lives, people often behave the way they typically do (they enact trait-congruent personality states), but occasionally behave differently from how they typically do (trait-incongruent personality states). Several theories propose that such incongruent personality states should be associated with undesirable outcomes such as less positive affect or more tiredness, but the current state of evidence is inconclusive and mostly based on one personality dimension: extraversion. In this study, we contribute to filling important gaps in the literature by examining congruence of personality dimensions other than extraversion, considering characteristics of the situation, and modeling congruence with state-of-the-art response surface analyses. We aimed to manipulate state honesty-humility and state agreeableness as well as perceived adversity and deception of the situation in a prisoner’s dilemma paradigm. The manipulations mostly had the intended effects but they also had additional unspecific effects on other personality states and situation characteristics. The study thus emphasized the difficulty of manipulating personality states, situation characteristics, and trait–state and state–situation congruence. In pre-registered analyses of variance, response surface analyses, and specification curve analyses, we then examined how trait–state congruence and state–situation congruence were associated with positive affect, tiredness, and performance in a numerical Stroop task. Neither trait–state congruence nor state–situation congruence were associated with positive affect, tiredness, or cognitive performance. However, in light of this study’s limitations, more studies that are carefully designed, carefully operationalized, and well-powered are needed to examine trait–state and state–situation congruence. Because experimental research can advance the understanding of personality dynamics substantially, future research should additionally further aim to develop valid and reliable manipulations of personality states and situation characteristics.

Personality traits describe how people typically think, feel, and behave (Diener & Lucas, 2020). However, the way people think, feel, and behave in a given moment—their personality states—can change from moment to moment and from situation to situation. People most often enact personality states at or close to their trait level (Fleeson, 2001). These personality states are called trait-congruent personality states. However, because people enact personality states across the whole spectrum of a given trait on a regular basis, their personality states occasionally differ from their trait level (Fleeson, 2001). These personality states are called trait-incongruent personality states. For example, people who are characterized by high levels on the personality trait extraversion nevertheless occasionally spend time alone (introverted personality state), people who are characterized by high levels on the personality trait agreeableness occasionally argue with others (disagreeable personality state), and people who are characterized by low levels on the personality trait conscientiousness occasionally complete their to-do list (conscientious personality state).

But does it matter whether a personality state is congruent or incongruent with one’s trait? Several theories (e.g., Gallagher et al., 2011; Moskowitz & Côté, 1995; Tett & Burnett, 2005) propose that people should experience more positive affect, less negative affect, and less tiredness when enacting trait-congruent personality states as opposed to...
trait-incongruent personality states. However, the current evidence regarding this question is rather inconclusive and previous research on the congruence of personality states has several gaps: First, almost all studies focused on extraversion (e.g., Fleeson et al., 2002; McNiel & Fleeson, 2006; Zelenski et al., 2012), whereas little is known about the other personality dimensions. Second, previous studies have examined a variety of outcomes of trait-incongruent personality states including positive affect, negative affect, effort, vitality, authenticity, performance in cognitive tasks, and handgrip strength, making it difficult to compare the findings (Fleeson & Wilt, 2010; Gallagher et al., 2011; McNiel et al., 2010; Pickett, Hofmans, Feldt, et al., 2020; Zelenski et al., 2012). Third, most studies modeled interactions between personality traits and personality states to examine whether the association between extraverted personality states and positive affect, for example, depends on the trait extraversion level (e.g., Gallagher et al., 2011; Jacques-Hamilton et al., 2018; Zelenski et al., 2012). However, interactions may not be the most adequate operationalization of congruence (J. R. Edwards, 2002; Humberg, Nestler, et al., 2019). Finally, although there is initial evidence that the situational context is relevant to the association between congruence and relevant outcomes (Gallagher et al., 2011; Whelan, 2013), characteristics of the situation were usually not considered.

The present study contributes to filling these gaps by (1) examining congruence of trait and state agreeableness and trait and state honesty-humility, (2) including and comparing three different outcomes, (3) modeling congruence with state-of-the-art response surface analyses (Cronbach, 1958; J. R. Edwards, 2002, 2007; Humberg, Nestler, et al., 2019), and (4) considering characteristics of the situation.

Consequences of Trait-Incongruent Personality States

Different theories have argued that trait-congruent personality states should be associated with more desirable outcomes than trait-incongruent personality states. The behavioral concordance model (Moskowitz & Côté, 1995) and trait activation theory (Tett & Burnett, 2005; Tett & Guterman, 2000), for example, propose that people experience more pleasant affect when engaging in trait-congruent personality states than in trait-incongruent states. Moreover, the contra-trait effort hypothesis (Gallagher et al., 2011) argues that incongruent personality states require more effort than congruent personality states. Therefore, trait-incongruent personality states should be associated with tiredness, fatigue, and even with impaired performance on subsequent tasks requiring similar effort (Gallagher et al., 2011; Zelenski et al., 2012). Finally, the trait-consistency hypothesis (Fleeson & Wilt, 2010) proposes that people feel more authentic when engaging in trait-congruent personality states than in trait-incongruent states.

Despite this theoretical basis, empirical findings on the consequences of trait-congruent and trait-incongruent personality states are mixed. Most prominently, a series of lab experiments examined associations between trait extraversion, state extraversion, and affect (e.g., Fleeson et al., 2002): Participants were instructed to be either talkative and bold (i.e., extraverted) or reserved and quiet (i.e., introverted) in a 10-minute group discussion. Afterwards, they rated their behavior and affect during the discussion. These experiments modeled interactions between personality traits and personality states to examine whether the association between personality states and positive affect depends on the personality trait level. However, they did not find any evidence for such an interaction; instead, all participants, regardless of their trait extraversion levels, reported higher levels of positive affect and lower levels of negative affect when behaving extravertedly (Fleeson et al., 2002; McNiel et al., 2010; McNiel & Fleeson, 2006; Zelenski et al., 2012; for a similar finding in an experience sampling study, see Margolis & Lyubomirsky, 2020). Similar findings were also reported regarding authenticity. People generally felt more authentic when their behavior was more open, conscientious, extraverted, agreeable, and emotionally stable, regardless of whether this behavior was congruent or incongruent with their personality trait (Cooper et al., 2018; e.g., Fleeson & Wilt, 2010; Wilt et al., 2021).

Other studies—mostly using experience sampling designs—did find evidence for trait-congruence effects. In an experimental experience sampling study, people low on trait extraversion reported weaker increases in positive affect and stronger increases in negative affect, inauthenticity, and tiredness when behaving extravertedly than people high on trait extraversion (Jacques-Hamilton et al., 2018). Moreover, people who behaved less extravertedly than normal in a group discussion reported more effort (Gallagher et al., 2011) and reacted more slowly in a subsequent Stroop task (Zelenski et al., 2012). Additionally, two recent experience sampling studies modeled congruence with quadratic regressions to examine whether the association between personality states and affect takes the form of an inverted U with its peak at the trait level. These studies found that people who behaved less conscientiously or more extravertedly than normal experienced less positive affect and less vitality, respectively, than people behaving congruently (Pickett, Hofmans, Debusscher, et al., 2020; Pickett, Hofmans, Feldt, et al., 2020). However, most of these effects were asymmetric such that, for example, only incongruent personality states below the trait extraversion level, but not above the trait level, were reported to be more effortful (Gallagher et al., 2011; Pickett, Hofmans, Debusscher, et al., 2020; Pickett, Hofmans, Feldt, et al., 2020). The associations between trait–state congruence and relevant outcomes may thus depend on the trait level and the direction of the incongruence.

These examples demonstrate that congruence has been studied with a variety of study designs (from laboratory experiments to longitudinal experience sampling designs and experimental experience sampling) and statistical operationalizations (from trait–state interactions to difference scores to quadratic associations). Particularly in the case of statistical operationalizations, it is important to consider how these operationalizations differ and under which circumstances they support the congruence hypothesis. A method that is well suited for this purpose is response surface analysis (Cronbach, 1958; J. R. Edwards, 2002, 2007; Humberg, Nestler, et al., 2019). Response surface analysis models congruence with polynomial regressions that in-
include linear effects, quadratic effects, and interactions. A significant congruence effect (typically called fit pattern) is not represented by a significant interaction coefficient or a significant quadratic coefficient as used in previous studies, but by a more complex pattern of associations involving both interactions and curvilinear effects (J. R. Edwards, 2007; Humberg, Dufner, et al., 2019). This approach is currently considered best practice for examining congruence effects (J. R. Edwards, 2007; Humberg, Dufner, et al., 2019). Moreover, response surface analysis has two important advantages: First, it includes interactions and quadratic effects and therefore makes it possible to compare the results to those found in previous studies. Second, response surface analysis can also model and detect asymmetric effects as reported in some of the previous studies.

Overall, we take away three key points: First, most previous research indicated that there is no interaction between personality traits and personality states with regard to positive affect, though a few recent studies suggest otherwise. Second, it may be important to examine outcomes other than positive affect such as tiredness, authenticity, or cognitive measures. Third, more complex statistical models such as response surface analyses could improve the understanding of the associations between trait-incongruent personality states and relevant outcomes.

**Trait–State Congruence in Context**

Another important limitation of previous research on trait–state congruence is that most previous studies failed to acknowledge possible influences of the situation. However, personality and behaviors are always contextualized in situations, and situations may therefore affect associations between trait–state congruence and relevant outcomes. Situations contain, for example, information such as incentives or rules about which behaviors are appropriate and will be rewarded and which behaviors are inappropriate and will be punished (Funder, 2016). Behaving extravertedly at a funeral, for example, may have completely different consequences than behaving extravertedly at a party.

Many of the previous studies examining trait–state congruence used a group discussion paradigm (e.g., Fleeson et al., 2002; McNiel et al., 2010; McNiel & Fleeson, 2006; Zeleinski et al., 2012). A group discussion represents a situation that probably calls for rather extraverted personality states. Extraverted personality states may be congruent with such a situation and therefore lead to more positive affect. In these studies, the situational context may therefore have confounded the results regarding trait–state congruence. Put differently: The reason why both people low and people high on trait extraversion experienced more positive affect when behaving extravertedly in the group discussions may be that extraverted personality states were appropriate to this situation. The associations between trait–state congruence and positive affect might be different when examined in a situation that does not reward extraverted personality states. Indeed, Whelan (2013) found that the association between extraverted personality states and positive affect weakened when the group discussion paradigm was changed to a group puzzling task. Moreover, Kritzler et al. (2020) found that the association between personality states and state affect was moderated by situation characteristics. For example, the association between state extraversion and state positive affect was stronger when the situation was perceived as more social.

Accordingly, theories such as person–environment fit (J. R. Edwards et al., 1998, 2006; Kristof, 1996) and trait-activation theory (Tett & Burnett, 2003; Tett & Guterman, 2000) propose that state–situation congruence, congruence between personality states and characteristics of the situation, is also associated with outcomes such as positive affect. That is, situation-congruent personality states should have positive consequences, whereas situation-incongruent personality states should have negative consequences. In the context of congruence with personality states, situations may best be conceptualized as situation characteristics from taxonomies such as the Situational Eight DIAMONDS (Rauthmann et al., 2014). Situation characteristics represent psychologically relevant and meaningful interpretations of situations such as their sociality, positivity, or adversity. Because these situation characteristics describe situations similar to the manner in which people can be described with trait (Rauthmann et al., 2014, p. 679) and have already been shown to be theoretically and empirically linked to personality dimensions such as the Big Five (e.g., Rauthmann et al., 2014, 2015, 2016), they are well suited for examining state–situation congruence.

Overall, state–situation congruence may have confounded the effects of trait–state congruence. Therefore, it is important to include state–situation congruence when examining trait–state congruence.

**An Experimental Situation for Manipulating Personality States and Situation Characteristics**

Like many previous studies, we aimed to examine trait–state congruence and state–situation congruence in an experimental study. An experimental manipulation of situation and behavior was necessary to ensure that sufficient congruent and incongruent personality states could be sampled. Moreover, experimental designs have the advantage that they allow causal inferences. However, a group discussion as used in most previous studies (e.g., Fleeson et al., 2002; McNiel & Fleeson, 2006) may not be well suited for our purposes because it is a situation that is difficult to control and manipulate.

We therefore sought a situation that is easier to control and manipulate and that is relevant to personality traits and personality states. We found that economic games such as the prisoner's dilemma (Flood, 1958; Rapoport, 1989; Tucker, 1980) are well-suited for this purpose because they have been associated with personality, and aspects of the situation have been previously manipulated (Kagel & McGee, 2014; Pothos et al., 2011; Zettler et al., 2013). In the prisoner's dilemma, two criminals are interrogated separately by the police for a crime. If both criminals remain silent (i.e., cooperate), they can only be convicted of a lesser crime and spend one year in prison. If one confesses and accuses the other (i.e., defects), the first one goes free while the other spends three years in prison. If both confess, they both go to prison for two years. The criminals have no way of knowing what the other will do. The dilemma in this sit-
ution is that each criminal is always better off by defecting than by cooperating, but that mutual cooperation yields a better outcome than mutual defection. In research, the prisoner’s dilemma is often reframed to have monetary rewards instead of prison sentences as payoffs (e.g., Kagel & McGee, 2014; Pothos et al., 2011).

Decisions in the prisoner’s dilemma are associated with the personality dimensions agreeableness (being patient, gentle vs. being quarrelsome, stubborn) and honesty-humility (being faithful, loyal vs. being sly, hypocritical) (Kagel & McGee, 2014; Pothos et al., 2011; Zettler et al., 2013). Moreover, both the prisoner’s dilemma and the personality dimensions agreeableness and honesty-humility are associated with the situation characteristics adversity (the extent to which the situation contains threats, problems, conflicts, competition, or criticism) and deception (the extent to which the situation contains mistrust, lying, or betrayal) (Rauthmann et al., 2014, 2015, 2016).

In our study, we aimed to manipulate state agreeableness and state honesty-humility as well as perceived adversity and deception in the prisoner’s dilemma. We manipulated state agreeableness and state honesty-humility together¹ by instructing participants to act either agreeably and honestly or disagreeably and dishonestly. We manipulated perceived adversity and deception together by manipulating the game’s payoffs and the description and behavior of the other player. We then examined how trait–state congruence and state–situation congruence in the prisoner’s dilemma were associated with relevant outcomes. Because trait–state congruence and state–situation congruence have been theoretically and empirically linked to a variety of outcomes (e.g., Gallagher et al., 2011; McNiel & Fleeson, 2006; Zelenski et al., 2012), we examined three different outcomes in our study: positive affect, tiredness, and cognitive performance.

Hypotheses

First, we expected the manipulation to be successful (H1). That is, we hypothesized that participants instructed to behave honestly and agreeably would be more honest, share more often, and report higher levels of state agreeableness and state honesty-humility than participants instructed to behave dishonestly and disagreeably. Moreover, we hypothesized that participants in the more adverse and deceptive condition would report higher levels of adversity and deception than participants in the less adverse and deceptive condition. In addition, we expected the experimental conditions to not be significantly associated with state emotionality, state extraversion, state conscientiousness, and state openness, nor with the situation characteristics duty, intellect, mating, positivity, negativity, and sociality. Second, because main effects of personality states and situation characteristics on state affect are well documented in the literature regardless of trait–state congruence and state–situation congruence (e.g., Ching et al., 2014; Kritzler et al., 2020; Sherman et al., 2015), we expected to replicate these main effects (H2). That is, we expected higher levels of state agreeableness and state honesty-humility to be associated with more positive affect, and higher levels of perceived adversity and deception to be associated with more negative affect. Finally, and most importantly, we expected trait-congruent personality states (H3) and situation-congruent personality states (H4) to be related to more positive affect, less tiredness, and better cognitive performance than trait-incongruent and situation-incongruent states.

Methods

Data collection was approved by the local ethics committee of the Ruhr University Bochum (Protocol 544: “Consequences of incongruent behaviors”).

The preregistration of sample sizes, hypotheses, and primary analyses is available on the Open Science Framework (Link: https://osf.io/hnu4b/). A summary of deviations from this preregistration, as well as study materials, data, analysis scripts, and supplementary materials are also available on the Open Science Framework (Link: https://osf.io/8tbeq/). We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study (Simmons et al., 2012).

Participants

We determined a required sample size of 152 participants based on an interaction effect of η² = .05 reported by Zelenski et al. (2012). However, because we suspected that effects might be smaller in our study, we aimed for a sample size of 200 participants with complete and valid data.

Participants were recruited via CloudResearch’s Prime Panels (Chandler et al., 2019). In total, 917 participants landed on the study entry page but 610 were excluded because they did not consent to participating in the study (N = 380), did not fulfill eligibility criteria (N = 35), did not pass the screening (N = 37), answered instructed response items incorrectly (N = 49), or dropped out early in the study. Figure 1 details the flow of participants through the study as well as exclusion criteria and final sample sizes. Because some participants terminated the study before completion (96 participants dropped out before the Stroop task)¹ and we preregistered using pairwise deletion, the sample sizes for the analyses differed depending on the measures included. Roughly speaking, we had a final sample of about 200 participants for analyses with Stroop performance as

¹ We manipulated state agreeableness and state honesty-humility, as well as adversity and deception because the two dimensions, respectively, are most likely highly correlated in the prisoner’s dilemma and it would thus have been difficult and unrealistic to manipulate them separately (Ashton et al., 2014; Rauthmann & Sherman, 2016a).
² Exploratory analyses of predictors of attrition revealed a significant correlation between progress in the study and trait extraversion, r = .22, p < .001, but not with any other personality variables, demographic variables, nor with the experimental conditions. 

Collabra: Psychology
dependent variable and of about 300 participants for analyses with positive affect or tiredness as dependent variable. Specifically, for analyses involving positive affect and tiredness as dependent variables, the final sample consisted of 307 participants. The participants were on average 56.82 years old (SD = 16.68, Q₁ = 44.00, Q₃ = 65.00, Q₅ = 69.00) and were evenly distributed between the genders with 154 male participants and 153 female participants. All participants lived in the United States and the majority (95.11%) spoke English as their first language.

Materials

Experimental Paradigm

Our experimental paradigm was a modified version of the classical prisoner’s dilemma (Flood, 1958; Tucker, 1980, 1983). We reframed the game to be about a jackpot of coins instead of prison sentences. Participants played 15 rounds against a computer-simulated player. In every round, the players had the opportunity to distribute a jackpot between them. Each player independently decided to either split the jackpot or steal the jackpot. If both chose to split, the jackpot was split between them. If one chose to steal while the other chose to split, the former received to jackpot whereas the latter received nothing. If both chose to steal, the jackpot was lost and both received nothing. To allow for more variation in honest and agreeable behavior, the players communicated to each other what they were planning to do before making their final decision—but they could lie or change their mind. We measured the number of honest decisions (i.e., trials in which the participants’ communicated intention and the final decision matched) and the number of share decisions (i.e., trials in which the participants’ final decision was to share the jackpot) as indices of participants’ behaviors during the game.

We aimed to manipulate behavior and situation independently in a 2 (behavior condition: high agreeableness and honesty vs. low agreeableness and honesty) × 2 (situation condition: high adversity and deception vs. low adversity and deception) between-person design. Behavior during the game was manipulated by directly instructing participants to act either honestly and agreeably or dishonestly and disagreeably. The situation was manipulated by varying payoff and information about the other player. In the low adversity and deception condition, payoffs signaled that stealing is not worthwhile (Zettler et al., 2013) and the other player was framed as and behaved like an trustworthy opponent. We tested an earlier version of the experimental paradigm in a preliminary study. Details on this pretest can be found in the Online Supplementary Materials (OSM).

Personality Traits

Personality traits were operationalized as self-reported HEXACO personality states (Ashton et al., 2004; Ashton & Lee, 2001) measured with the 24-item Brief HEXACO Inventory (De Vries, 2013). Participants indicated on a scale from 1 (disagree strongly) to 5 (agree strongly) how much they agreed with descriptive statements such as “I think science is boring,” an item for Openness. Scale scores were computed by reverse-coding items when applicable and then averaging over all items that belonged to a scale. The personality trait scales showed rather low internal consistencies (Honesty-Humility: α = .61, Emotionality: α = .28, eXtraversion: α = .59; Agreeableness: α = .22, Conscientiousness: α = .52, Openness: α = .54) which is to be expected for short scales that are designed to maximize coverage of the construct and convergent validity instead of internal consistency (De Vries, 2013; Ziegler et al., 2014). However, the internal consistencies for Agreeableness and Emotionality were even lower than expected. This limitation is further addressed in the discussion.

Personality States

Personality states were operationalized as self-reported HEXACO personality states measured with bipolar items adopted from Sherman et al. (2015) and Churchyard (2015). All personality states were measured with one bipolar item each adopted from Sherman et al. (2015). State honesty-humility and state agreeableness were additionally measured with three items from Churchyard (2013), and all four items were combined into one scale score for these states. Participants indicated on a 7-point bipolar scale how they behaved in the game situation (e.g., for one of the agreeableness items, the scale ranged from cold, quarrelsome to warm, agreeable). Scale scores (when applicable) were computed by reverse-coding items when applicable and then averaging over all items that belonged to a scale.

Situation Characteristics

Situation characteristics were operationalized as per-

3 The exact wording of the instructions was: “Your task is to behave honestly and cooperatively with your [partner/opponent]. Please try to be true to your word most of the time. Of course, you do not always have to do as you proposed, but please try to be honest and cooperative as much as you can.” in the high agreeableness and honesty-condition and “Your task is to behave dishonestly and unco-operatively with your [partner/opponent]. Please try to play dirty most of the time. Of course, you do not always have to deceive your [partner/opponent], but please try to be dishonest and uncooperative as much as you can.” in the low agreeableness and honesty-condition.
ceived situation characteristics from the Situational Eight DIAMONDS taxonomy (Rauthmann et al., 2014). The situation characteristics adversity and deception were measured with three items each taken from the S8* questionnaire (Rauthmann & Sherman, 2016a). The other situation characteristics were measured with one item each taken from the S8-I questionnaire (Rauthmann & Sherman, 2016b). Participants indicated on a scale from 1 (not at all) to 7 (totally) how well certain statements described the game situation they had just encountered (e.g., for deception, one item was “It is possible to deceive someone.”). Scale scores (when applicable) were computed by reverse-coding items when applicable and then averaging over all items that belonged to a scale.

**Positive Affect and Tiredness**

Positive affect was operationalized as the good–bad dimension and tiredness was operationalized as the active–tired dimension of the three-dimensional model of affect (e.g., Matthews et al., 1990; Steyer et al., 1994; Wilhelm & Schoebl, 2007). These dimensions assess the valence (from unpleasant to pleasant) and the energetic arousal (from tired/without energy to awake/full of energy) of the current mood. Positive affect and tiredness were measured with a short form of the multidimensional mood state questionnaire (Steyer, 2021; Steyer et al., 1994) comprising four items each. Participants indicated on a scale from 1 (not at all) to 7 (totally) how intensely they had experienced moods characterized by certain adjectives at the end of the game (e.g., “content” and “tired”). Scale scores were computed by reverse-coding items when applicable and then averaging over all items that belonged to a scale.

**Stroop Task**

Participants performed a numerical Stroop task (Henik & Tzelgov, 1982) as a measure of cognitive performance. In this task, participants were presented two numbers (ranging from 1 to 9) and indicated which number is numerically larger. The two numbers always had a numerical difference of one and were presented in different physical sizes: In neutral trials, both numbers had the same physical size and only differed in their numerical size. In congruent trials,
the numerically larger number was also displayed physically larger. In incongruent trials, the numerically smaller number was displayed physically larger.

Participants first completed a trial round with immediate feedback. The main task then consisted of 75 trials with 15 neutral, 30 congruent, and 30 incongruent trials presented in a random order. Participants were instructed to answer as fast as possible but correctly. We sampled three types of information during the main Stroop task: reaction time, trial type (congruent, incongruent, or neutral), and correctness of the answer.

**Comprehension, Attention, and Careless Responding**

We included a short screening consisting of four synonym tasks (see Chandler et al., 2019) at the beginning of the study to assess language comprehension and attentiveness. Additionally, we included two instructed response items (e.g., ”To secure data quality, please choose answer option 4”; Meade & Craig, 2012) to further identify careless responders. Only participants who passed the screening and provided correct responses on the instructed response items were included in the final sample.

**Procedure**

The study was implemented using the online survey platform Qualtrics (Qualtrics, 2020). To be eligible, participants had to be at least 18 years old, live in the United States, access the study with a laptop or desktop computer, and consent to participate. After providing informed consent, the participants completed the screening task. Participants who did not fulfill screening criteria or did not answer the screening questions correctly were screened out and did not complete the rest of the study.

Participants who passed the screening then answered questions regarding their personality traits and demographics. They then played the prisoner’s dilemma game which took on average 14 minutes (Q₁ = 5.65, Q₂ = 7, Q₃ = 8.19). Next, participants answered questions regarding their personality states, perceived situation characteristics, and their positive affect and tiredness at the end of the game in randomized order. Finally, participants performed the numerical Stroop task (duration ~ 7 minutes). In the end, participants were given the chance to comment on the study, created their individual code for further reference, and were compensated for study participation. Importantly, participants’ compensations did not depend on the results of the prisoner’s dilemma game. Overall, participants took on average 33 minutes (median = 24, SD = 85) to complete the study.

**Analytic Strategy**

Analyses were conducted in R (Version 4.0.3; R Core Team, 2020) with the R packages tidyverse (Version 1.5.0; Wickham et al., 2019), afex (Version 0.28.1; Singmann et al., 2021), RSA (Version 0.10.2; Schönbrodt & Humberg, 2021), specr (Version 0.2.1; Masur & Scharkwok, 2019), and papaja (Version 0.1.0.9997; Aust & Barth, 2020). Detailed information about the planned analysis strategy can be found in the preregistration. Here, we only report on the analyses that were ultimately realized.

For the manipulation check (H1), we conducted analyses of variance with the different situation characteristics and personality states as dependent variables and effects-coded variables indicating the experimental situation and behavior conditions and their interaction as independent variables. The main effects of personality states and situation characteristics on the positive affect dimension (H2) were examined in a multiple regression analysis. The planned analyses for the congruence hypotheses (H3 and H4) depended on the results from the manipulation check. In short, we planned to perform analyses of variance with the experimental groups and median-split personality traits whenever a manipulation was successful. In addition, and regardless of the success of the manipulations, we conducted response surface analyses (based on polynomial regression) with self-reported situation perceptions, personality states, and personality traits. Moreover, we conducted specification curve analyses for all analyses with cognitive performance indicators as dependent variables. Details on response surfaces analysis and specification curve analysis can be found in the respective sections. Because participants were nested within experimental groups, we controlled for these groups in all analyses regarding congruence hypotheses. We additionally estimated models controlling for the results of the game (i.e., points won and difference between participants’ points and the computer’s points).

Overall, we examined two combinations for trait-state congruence (i.e., trait agreeableness–state agreeableness and trait honesty-humility–state honesty-humility) and four combinations for state–situation congruence (i.e., state agreeableness-adversity, state agreeableness–deception, state honesty-humility–adversity, and state honesty-humility–deception). Moreover, we considered two different operationalizations of congruence (i.e., interactions to

---

4 We, furthermore, used the R-packages apaTables (Version 2.0.8; Stanley, 2021), BayesFactor (Version 0.9.12.4.2; Morey & Rouder, 2018), car (Version 5.0.10; Fox et al., 2020; Fox & Weisberg, 2019), carData (Version 3.0.4; Fox et al., 2020), circularize (Version 0.4.12; Gu et al., 2014), coda (Version 0.19.4; Plummer et al., 2006), cowplot (Version 1.1.1; Wilke, 2020), dplyr (Version 1.0.6; Wickham et al., 2021), forcats (Version 0.5.1; Wickham, 2021a), ggplot2 (Version 3.5.4; Wickham, 2016), inlcar-package, Fisher, 2020, irr (Version 0.84.1; Gamer et al., 2019), jpeg (Version 0.1.8.1; Urbanek, 2019), knitr (Version 1.33; Xie, 2015), lattice (Version 0.20.41; Sarkar, 2008), lavaan (Version 0.6.8; Rosseel, 2012), lemon (Version 0.4.5; S. M. Edwards, 2020), lme4 (Version 1.1.27; Bates et al., 2015), lpSolve (Version 5.6.15; Bergelkaer, 2020), Matrix (Version 1.2.18; Bates & Maechler, 2019), patchwork (Version 1.1.1; Pedersen, 2020), psych (Version 2.1.5; Revelle, 2021), purrr (Version 0.3.4; Henry & Wickham, 2020), readr (Version 1.4.0; Wickham & Hester, 2020), readxl (Version 1.3.1; Wickham & Bryan, 2019), stringr (Version 1.4.0; Wickham, 2019), tidytext (Version 1.3.2; Müller & Wickham, 2021), tidyr (Version 1.1.3; Wickham, 2021b), tinylabels (Version 0.2.1; Barth, 2021), and yarr (Version 0.1.5; Phillips, 2017).
compare the results to previous studies and fit patterns to use the best practice response surface analyses) and four different outcomes (i.e., positive affect, tiredness, Stroop reaction times, and Stroop error rates). We thus examined a total of 6 (congruence) × 2 (congruence operationalization) × 4 (outcomes) = 48 associations between congruence and relevant outcomes. Because a maximum of eight of these were used to test the same hypothesis, we used a Bonferroni-corrected significance level of α = .00625 for these analyses.

Response Surface Analysis

A congruence effect is typically defined as a pattern in which there are certain combinations of two predictors (e.g., equal levels of personality trait and personality state) that are associated with optimal values of an outcome whereas any deviations from these optimal combinations are associated with decreases in the outcome (Humberg, Nestler, et al., 2019; e.g., Schönbrodt, 2016). As discussed above, trait–state congruence has been examined with a variety of statistical operationalizations including interactions, difference scores, and quadratic associations. Importantly, each of these may be indicative of congruence effects, but only under certain circumstances. For example, a significant interaction does not necessarily indicate a congruence effect, but an interaction that causes the direction of the relationship to change (e.g., Jacques-Hamilton et al., 2018) could indicate a congruence effect (compare Figures 2A and 2B). Similarly, a significant negative quadratic association does not necessarily indicate a congruence effect (e.g., Cooper et al., 2018), but a significant negative quadratic association with a peak near the trait level does (compare Figures 2C and 2D). Thus, interactions and quadratic associations may be indicative of congruence effects but only under certain conditions. Response surface analysis—which includes interactions and quadratic associations—facilitates the testing of congruence effects because it precisely defines these conditions. In response surface analyses, a congruence effect is represented by a specific and more complex pattern of associations called a fit pattern (see Figure 2E).

Response surface analyses (Cronbach, 1958; J. R. Edwards, 2002, 2007; Humberg, Nestler, et al., 2019) are based on polynomial regressions that include main effects, squared effects, and interactions. Parameters of the model are then statistically and graphically interpreted, and the predicted values are plotted as a three-dimensional surface to facilitate their interpretation. A significant congruence effect is represented by a specific pattern of associations in a polynomial regression (J. R. Edwards, 2007; see Humberg, Dufner, et al., 2019). This fit pattern is defined by specific parameters of the response surface which test the shape of the surface above the line of congruence (i.e., where X = Y) and the line of incongruence (i.e., where X = -Y) and examine the position of the ridge of the surface (the specific criteria for a fit pattern in a broad sense are a0 < 0, a2 = 0, a3 = 0, and a2 = 0; see Humberg, Nestler, et al., 2019; Nestler et al., 2019). A significant fit pattern indicates that the highest values of a dependent variable (e.g., positive affect) are predicted when the values of both predictor variables are equal, and that the predicted values decrease the more the two predictor variables deviate from each other. As mentioned before, we used both interactions and fit patterns as operationalizations of congruence in this study. Figures 2E and 2F compare response surfaces for congruence operationalized as an interaction (2E) and as a fit pattern (2F): The two approaches are similar (in fact, interactions are a part of the fit pattern), but still differ meaningfully.

For the response surface analyses, personality traits were transformed from a 5-point scale to a 7-point scale. Personality traits, personality states, and situation characteristics were centered on the midpoints of their respective scales (J. R. Edwards & Parry, 1995; Humberg, Nestler, et al., 2019).

Specification Curve Analysis

We conducted specification-curve analyses (Simonsohn et al., 2020) to account for the large number of researcher degrees of freedom in identifying and handling outliers and calculating the summary statistic in the Stroop task (Simmons et al., 2011). The specification curve analyses involved three steps. First, we determined the set of reasonable specifications. We specified different choices for identifying outliers, handling these outliers, selecting the trials to be considered for the calculation of the summary statistic, calculating the summary statistic, and including covariates. In sum, we analyzed a total of 8896 reaction time specifications and 2224 error rate specifications. Details on the set of reasonable specifications can be found in the OSM.

In the next step, the analyses were conducted separately with each specification as the dependent variable and the results were graphically illustrated in descriptive specification curves. These curves displayed the effects of interest (i.e., interaction coefficients or presence of fit patterns) from all specifications ranked by size of the effect. Additionally, these plots indicate which analytic decisions lead to which effects. Descriptive specification curves were used to examine the range of the effects across all specifications, the proportion of significant effects, and how the analytic decisions impacted the estimated effects.

Finally, to come to an overall conclusion regarding the presence of the hypothesized effect, we applied a permutation technique to test how inconsistent the results were with the null hypothesis of no effect across the specification curve. We created 500 data sets by shuffling the values within the dependent variables (i.e., the different reaction time and error rate specifications) and repeated the response surface analyses with these shuffled data. Because of the random distribution of reaction times or error rates, the null hypothesis of no effect is by definition true in these shuffled datasets. The number of shuffled data with at least as many significant specifications as the original data was then used to calculate a p value of the permutation test. Additionally, we graphically compared the specification curve of the effects obtained in the original data to the range of the expected-under-the-null specification curves from the shuffled datasets.
Figure 2. Exemplary response surface plots of congruence operationalized as an interaction (A) and congruence operationalized as a fit pattern as defined in response surface analyses (B).

Note. Association between two predictor variables X and Y and an outcome variable Z are plotted as a three-dimensional surface. The color of the surface indicates the predicted levels of the outcome variable with greener colors indicating higher levels and orange to red colors indicating lower levels. The blue lines represent the line of congruence (LOC; i.e., X = Y) and the line of incongruence (LOIC; i.e., X = -Y).

Results

Table 1 displays means, standard deviations, and bivariate correlations of the relevant study variables. Because of the high number of different analyses, not all results can be shown and discussed in detail in the main manuscript. Here, we focus on the overarching pattern of results. The online supplementary materials (OSM) entail additional results, tables, and figures.

Manipulation Check

As expected, participants in the high agreeableness and honesty condition reported significantly higher levels of state agreeableness ($M = 5.23$, $SD = 1.40$) and state honesty-humility ($M = 5.30$, $SD = 1.51$) than participants in the low agreeableness and honesty condition ($M_{\text{agree}} = 4.32$, $SD_{\text{agree}} = 1.65$, $d_{\text{agree}} = 0.60$, 95% CI [0.34; 0.85]; $M_{\text{honest}} = 5.58$, $SD_{\text{honest}} = 1.92$, $d_{\text{honest}} = 1.00$, 95% CI [0.72; 1.27]; full results in Table 2). Moreover, participants in the low adversity and deception condition perceived the game situation to have significantly lower levels of adversity ($M = 2.22$, $SD = 1.38$) but not deception ($M = 5.27$, $SD = 1.40$) than participants in the high adversity and deception condition ($M_{\text{adv}} = 2.65$, $SD_{\text{adv}} = 1.70$, $d_{\text{adv}} = 0.28$, 95% CI [0.06; 0.51]; $M_{\text{dec}} = 5.43$, $SD_{\text{dec}} = 1.43$, $d_{\text{dec}} = 0.11$, 95% CI [-0.11; 0.54]; full results in Table 2).
Table 1. Bivariate correlations, means, and standard deviations of the study variables

|                      | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   | 18   | 19   | 20   | 21   | 22   | 23   | M    | SD   |
|----------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Trait Honesty-Humility | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 4.04 | 0.77 |
| Trait Emotionality    | -.06 | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 2.88 | 0.66 |
| Trait Extraversion    | -.21 | -.19 | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 3.62 | 0.77 |
| Trait Agreeableness   | .08  | -.11 | .29  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 3.01 | 0.57 |
| Trait Conscientiousness | -.05 | -.12 | .15  | .08  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 3.37 | 0.51 |
| Trait Openness        | -.17 | -.12 | .19  | .19  | .25  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 3.38 | 0.72 |
| State Honesty-Humility | -.05 | -.01 | .15  | .23  | .14  | .09  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 4.43 | 1.93 |
| State Emotionality    | -.18 | .24  | -.11 | -.07 | -.02 | -.02 | -.14 | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 2.68 | 1.73 |
| State Extraversion    | -.24 | -.03 | .19  | .05  | .16  | .15  | .28  | .15  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 3.88 | 1.91 |
| State Agreeableness   | .00  | -.10 | .18  | .24  | .15  | .11  | .82  | -.24 | .34  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 4.77 | 1.60 |
| State Conscientiousness | -.03 | -.10 | .17  | .17  | .22  | .12  | .45  | -.18 | .25  | .53  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 5.22 | 1.64 |
| State Openness        | -.13 | -.14 | .10  | .21  | .20  | .24  | .41  | -.06 | .31  | .47  | .61  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 5.09 | 1.66 |
| Sharing Behaviors     | .17  | .02  | .03  | .08  | -.07 | -.09 | .50  | -.05 | .01  | .39  | .11  | -.02 | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 7.25 | 4.68 |
| Honest Behaviors      | .06  | .04  | -.01 | .05  | .03  | -.01 | .23  | .05  | .01  | .14  | .06  | .01  | .47  | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 10.7 | 3.63 |
| Duty                  | -.23 | .00  | .10  | .11  | .20  | .07  | .03  | .07  | .10  | .01  | .22  | .11  | -.10 | -.06 | -    | -    | -    | -    | -    | -    | -    | -    | -    | -    | 4.42 | 1.89 |
| Intellect             | -.25 | .05  | .08  | .24  | .17  | .23  | .16  | .12  | .24  | .18  | .21  | .25  | -.06 | -.12  | .35  | -    | -    | -    | -    | -    | -    | -    | -    | 3.69 | 2.02 |
| Adversity             | -.45 | .10  | -.20 | .04  | .14  | .06  | -.08 | .27  | .00  | -.19 | -.09 | .01  | -.15  | -.12  | .30  | .32  | -    | -    | -    | -    | -    | -    | -    | 2.42 | 1.56 |
| Mating                | -.54 | -.01 | -.01 | .08  | .12  | .16  | .20  | .18  | .30  | .15  | .07  | .16  | .01  | -.03  | .32  | .42  | .49  | -    | -    | -    | -    | -    | 2.30 | 1.79 |
| Positivity            | -.21 | -.11 | .08  | .19  | .08  | .09  | .52  | -.16 | .37  | .54  | .26  | .34  | .16  | .04  | .17  | .31  | .01  | .37  | -    | -    | -    | -    | -    | 3.90 | 1.88 |
| Negativity            | -.18 | .12  | -.11 | -.10 | .13  | .02  | -.30 | .23  | -.21 | -.34 | -.17 | -.16 | -.08 | -.04 | .18  | .15  | .55  | .17  | -.31 | -    | -    | -    | -    | -    | 3.55 | 2.11 |
| Deception             | .11  | .04  | .03  | -.14 | .01  | -.05 | -.16 | -.10 | -.16 | .06  | .08  | -.04 | -.04 | -.11 | .02  | .00  | -.10 | .15  | .17 | .27  | -    | -    | -    | -    | 5.35 | 1.41 |
| Sociality             | -.21 | .06  | .07  | .06  | .03  | .07  | .12  | .08  | .31  | .10  | .07  | .04  | -.03 | -.09  | .31  | .39  | .28  | .35  | .22 | .18  | .12  | -    | -    | 4.27 | 1.98 |
| Positive Affect       | -.01 | -.18 | .20  | .13  | .07  | .03  | .52  | -.26 | .28  | -.52 | .43  | .43  | .18  | .13  | .04  | .12  | -.27 | -.07 | .57  | -.53 | -.14 | .03  | -    | 3.59 | 1.05 |
| Tiredness(r)          | .01  | -.16 | .21  | .19  | .11  | .10  | .37  | -.19 | .30  | .41  | .34  | .38  | .12  | .16  | .06  | .10  | -.20 | .08  | .49  | -.37 | -.09 | .07  | .67  | 3.39 | 0.92 |

Note. * p < 0.05; ** p < 0.01; *** p < 0.001. (r) indicates that tiredness was reverse-coded such that higher values indicate less tiredness.
| Predictor | SS    | df | MS   | F     | p     | $\eta_p^2$ | 95% CI for $\eta_p^2$ |
|-----------|-------|----|------|-------|-------|-----------|------------------------|
| DV: Number of share decisions |       |    |      |       |       |           |                        |
| (Intercept) | 11500.36 | 1  | 11500.36 | 762.99 | < .001 | .16 | [.08, .24] |
| Behavior condition | 704.80 | 1  | 704.80 | 46.76 | < .001 | .18 | [.10, .27] |
| Situation condition | 827.35 | 1  | 827.35 | 54.89 | < .001 | .18 | [.10, .27] |
| Behavior condition x Situation condition | 86.15 | 1  | 86.15 | 5.72 | .018 | .02 | [.00, .07] |
| Error | 3707.91 | 246 | 15.07 |       |       |           |                        |
| DV: Number of honest decisions |       |    |      |       |       |           |                        |
| (Intercept) | 33960.13 | 1  | 33960.13 | 2715.74 | < .001 | .04 | [.01, .10] |
| Behavior condition | 161.21 | 1  | 161.21 | 12.89 | < .001 | .04 | [.01, .10] |
| Situation condition | 12.21 | 1  | 12.21 | 0.98 | .324 | .00 | [.00, .03] |
| Behavior condition x Situation condition | 53.31 | 1  | 53.31 | 4.26 | .040 | .01 | [.00, .05] |
| Error | 3663.94 | 293 | 12.50 |       |       |           |                        |
| DV: State agreeableness |       |    |      |       |       |           |                        |
| (Intercept) | 5815.05 | 1  | 5815.05 | 2505.20 | < .001 | .08 | [.03, .15] |
| Behavior condition | 50.88 | 1  | 50.88 | 21.92 | < .001 | .08 | [.03, .15] |
| Situation condition | 6.64 | 1  | 6.64 | 2.86 | .092 | .01 | [.00, .05] |
| Behavior condition x Situation condition | 6.12 | 1  | 6.12 | 2.64 | .106 | .01 | [.00, .05] |
| Error | 587.26 | 253 | 2.32 |       |       |           |                        |
| DV: State honesty-humility |       |    |      |       |       |           |                        |
| (Intercept) | 5018.31 | 1  | 5018.31 | 1685.00 | < .001 | .20 | [.12, .28] |
| Behavior condition | 188.10 | 1  | 188.10 | 63.16 | < .001 | .20 | [.12, .28] |
| Situation condition | 9.36 | 1  | 9.36 | 3.14 | .077 | .01 | [.00, .05] |
| Behavior condition x Situation condition | 1.26 | 1  | 1.26 | 0.42 | .516 | .00 | [.00, .03] |
| Error | 753.49 | 253 | 2.98 |       |       |           |                        |
| DV: Adversity |       |    |      |       |       |           |                        |
| (Intercept) | 1794.44 | 1  | 1794.44 | 754.97 | < .001 | .930 | [.00, .01] |
| Behavior condition | 0.02 | 1  | 0.02 | 0.01 | .930 | .00 | [.00, .01] |
| Situation condition | 15.31 | 1  | 15.31 | 6.44 | .012 | .02 | [.00, .06] |
| Behavior condition x Situation condition | 6.89 | 1  | 6.89 | 2.90 | .090 | .01 | [.00, .04] |
| Error | 713.05 | 300 | 2.38 |       |       |           |                        |
| DV: Deception |       |    |      |       |       |           |                        |
| (Intercept) | 8656.47 | 1  | 8656.47 | 4354.99 | < .001 | .147 | [.00, .04] |
| Behavior condition | 4.20 | 1  | 4.20 | 2.11 | .147 | .01 | [.00, .04] |
| Situation condition | 1.92 | 1  | 1.92 | 0.97 | .326 | .00 | [.00, .03] |
| Behavior condition x Situation condition | 1.74 | 1  | 1.74 | 0.88 | .350 | .00 | [.00, .03] |
| Error | 596.31 | 300 | 1.99 |       |       |           |                        |
| Predictor                        | SS       | df | MS      | F       | \(p\) | \(\eta_p^2\) | 95% CI for \(\eta_p^2\) |
|--------------------------------|--|-----|--------|--------|-------|-------------|------------------|
| **DV: State openness**          |          |     |        |        |       |             |                  |
| (Intercept)                     | 6628.26  | 1   | 6628.26| 2478.05| < .001| .03         | [.00, .08]       |
| Behavior condition              | 18.27    | 1   | 18.27  | 6.83   | .009  | .03         | [.00, .08]       |
| Situation condition             | 1.33     | 1   | 1.33   | 0.50   | .481  | .00         | [.00, .03]       |
| Behavior condition x Situation condition | 7.13    | 1   | 7.13   | 2.67   | .104  | .01         | [.00, .05]       |
| Error                           | 676.72   | 253 | 2.67   |        |       |             |                  |
| **DV: State conscientiousness** |          |     |        |        |       |             |                  |
| (Intercept)                     | 6994.47  | 1   | 6994.47| 2668.09| < .001| .03         | [.00, .08]       |
| Behavior condition              | 22.05    | 1   | 22.05  | 8.41   | .004  | .03         | [.00, .08]       |
| Situation condition             | 0.16     | 1   | 0.16   | 0.06   | .806  | .00         | [.00, .02]       |
| Behavior condition x Situation condition | 0.40    | 1   | 0.40   | 0.15   | .698  | .00         | [.00, .02]       |
| Error                           | 663.25   | 253 | 2.62   |        |       |             |                  |
| **DV: State emotionality**      |          |     |        |        |       |             |                  |
| (Intercept)                     | 1835.05  | 1   | 1835.05| 615.68 | < .001| .02         | [.00, .07]       |
| Behavior condition              | 15.20    | 1   | 15.20  | 5.10   | .025  | .02         | [.00, .07]       |
| Situation condition             | 0.15     | 1   | 0.15   | 0.05   | .822  | .00         | [.00, .02]       |
| Behavior condition x Situation condition | 0.00    | 1   | 0.00   | 0.00   | .973  | .00         | [.00, 1.00]      |
| Error                           | 754.08   | 253 | 2.98   |        |       |             |                  |
| **DV: Positivity**              |          |     |        |        |       |             |                  |
| (Intercept)                     | 4656.75  | 1   | 4656.75| 1366.85| < .001| .04         | [.01, .09]       |
| Behavior condition              | 41.05    | 1   | 41.05  | 12.05  | .001  | .04         | [.01, .09]       |
| Situation condition             | 11.05    | 1   | 11.05  | 3.24   | .073  | .01         | [.00, .04]       |
| Behavior condition x Situation condition | 0.13    | 1   | 0.13   | 0.04   | .848  | .00         | [.00, .01]       |
| Error                           | 1032.30  | 303 | 3.41   |        |       |             |                  |

*Note.* The lines separate the different models from each other.
However, contrary to our hypotheses, participants in the high agreeableness and honesty condition also reported significantly higher levels of state openness and perceived positivity and significantly lower levels of state emotionality than participants in the low agreeableness and honesty condition (see Table 2). Moreover, we found that participants’ concrete behaviors in the game (i.e., deciding to share and being honest when communicating the intended decision) were significantly associated with both the behavior conditions, the situation conditions, and with their interaction. Thus, the manipulations also had unspecific effects on top of the expected effects.

In sum, the manipulation had both specific effects and unspecific effects. Because the manipulation had the expected specific effects (albeit with small effect sizes in some cases), we decided to perform the planned analyses for the case of a successful manipulation. However, we would also like to explicitly point out that these results should be interpreted with caution due to the unspecific effects.

**Replicating Main Effects**

As expected, people who reported higher levels of state agreeableness, \( \beta = 0.25, 95\% CI [0.05, 0.44], t(226) = 2.50, p = .013 \), higher levels of state honesty-humility, \( \beta = 0.32, 95\% CI [0.10, 0.54], t(226) = 2.88, p = .004 \), and lower levels of adversity of the situation, \( \beta = -0.21, 95\% CI [-0.32, -0.09], t(226) = -3.59, p < .001 \), experienced more positive affect. In contrast, perceived deception was not significantly associated with positive affect, \( \beta = -0.03, 95\% CI [-0.15, 0.09], t(226) = -0.51, p = .614 \). Adding personality traits and interactions between personality traits and personality states as predictors did not change the pattern of associations meaningfully. Detailed results from these analyses can be found in the OSM.

**Associations Between Congruence and Positive Affect**

Analyses of variance predicting positive affect from interactions between behavior conditions and median-split personality traits and from interactions between behavior conditions and situation conditions did not yield any significant interaction effects (all \( F_s < 2.20, all \ p_s > .132 \)). Instead, only the behavior conditions were consistently associated with reverse-coded tiredness (all \( F_s > 5.84, all \ p_s < .016 \)) such that participants in the high agreeableness and honesty condition (\( M = 3.54, SD = 0.78 \)) reported significantly more active mood—that is, less tiredness—than participants in the low agreeableness and honesty condition (\( M = 3.25, SD = 1.02, d = 0.33, 95\% CI [0.08, 0.58] \)). Including results of the game as covariates did not change anything about this pattern of results. Detailed results from these analyses can be found in the OSM. Overall, the analyses of variance did not yield any evidence for trait-state or state-situation congruence effects with respect to tiredness.

Next, we conducted response surface analyses with self-reported personality traits, personality states, and situation perceptions and found a very similar pattern of results (Table 4): Higher levels of state agreeableness and state honesty-humility were associated with more active mood in all models except one (bs ranging from 0.15 to 0.26). In contrast, neither personality traits nor situation characteristics nor any of the trait-state interactions, state-situation interactions, or curvilinear coefficients were significantly associated with positive affect in any of the models (Table 3).

**Associations Between Congruence and Tiredness**

Analyses of variance predicting tiredness from interactions between behavior conditions and median-split personality traits and from interactions between behavior conditions and situation conditions did not yield any significant interaction effects (all \( F_s < 2.20, all \ p_s > .139 \)). Instead, only the behavior conditions were consistently associated with reverse-coded tiredness (all \( F_s > 5.84, all \ p_s < .016 \)) such that participants in the high agreeableness and honesty condition (\( M = 3.54, SD = 0.78 \)) reported significantly more active mood—that is, less tiredness—than participants in the low agreeableness and honesty condition (\( M = 3.25, SD = 1.02, d = 0.33, 95\% CI [0.08, 0.58] \)). Including results of the game as covariates did not change anything about this pattern of results. Detailed results from these analyses can be found in the OSM. Overall, the analyses of variance did not yield any evidence for trait-state or state-situation congruence effects with respect to tiredness.

Next, we conducted response surface analyses with self-reported personality traits, personality states, and situation perceptions and found a very similar pattern of results (Table 4): Higher levels of state agreeableness and state honesty-humility were associated with more active mood in all models except one (bs ranging from 0.15 to 0.26). In contrast, neither personality traits nor situation characteristics nor any of the trait-state interactions, state-situation interactions, or curvilinear coefficients were significantly associated with tiredness in any of the models (Table 4).

**Figure 3** shows the response surface plots for the associations between personality traits, personality states, and positive affect (Panels A and B) and between personality states, situation characteristics, and positive affect (Panels C to F). The surfaces were mostly steep along the personality-state axis. This illustrates the results from polynomial regressions that personality states were the only significant predictors of positive affect. Additionally, none of the response surfaces fulfilled criteria for a fit pattern (i.e., \( a_2 = 0, a_4 < 0, \text{and } a_6 = 0 \)). Thus, trait-state congruence and state-situation congruence were not significantly associated with positive affect.
Table 3. Results from polynomial regression models of association between positive affect and congruence

| Term                                      | $b$  | SE  | 95% CI          | $\beta$ | $p$  |
|-------------------------------------------|------|-----|-----------------|---------|------|
| **Trait–state congruence: Agreeableness** |      |     |                 |         |      |
| Intercept                                 | 0.04 | 0.05| [-0.05, 0.13]   | 0.12    | .404 |
| State A                                   | 0.28 | 0.05| [0.19, 0.38]    | 0.43    | < .001 |
| Trait A                                   | 0.01 | 0.08| [-0.15, 0.17]   | 0.01    | .906 |
| State $A^2$                               | 0.04 | 0.02| [0.00, 0.08]    | 0.14    | .035 |
| State A x Trait A                         | -0.01| 0.04| [-0.09, 0.06]   | -0.02   | .704 |
| Trait $A^2$                               | 0.01 | 0.04| [-0.07, 0.09]   | 0.01    | .808 |
| Behavior conditions                       | 0.19 | 0.12| [-0.04, 0.41]   | 0.09    | .107 |
| Situation conditions                      | -0.10| 0.12| [-0.32, 0.13]   | -0.05   | .405 |
| Points                                    | 3.10 | 0.28| [2.56, 3.64]    | 2.98    | < .001 |
| Point difference                          | 0.29 | 0.09| [0.12, 0.46]    | 0.44    | .001 |
| **Model Fit**                             |      |     |                 |         |      |
|                                           |      |     | $R^2 = .29$     |         |      |
| **Trait–state congruence: Honesty-Humility** |      |     |                 |         |      |
| Intercept                                 | 0.08 | 0.04| [-0.01, 0.17]   | 0.27    | .668 |
| State HH                                  | 0.20 | 0.06| [0.09, 0.32]    | 0.37    | .001 |
| Trait HH                                  | -0.13| 0.10| [-0.32, 0.06]   | -0.14   | .186 |
| State $HH^2$                              | 0.00 | 0.02| [-0.04, 0.03]   | -0.01   | .902 |
| State HH x Trait HH                       | 0.04 | 0.03| [-0.02, 0.10]   | 0.15    | .176 |
| Trait $HH^2$                              | 0.05 | 0.04| [-0.03, 0.12]   | 0.13    | .224 |
| Behavior conditions                       | 0.04 | 0.13| [-0.21, 0.29]   | 0.02    | .747 |
| Situation conditions                      | -0.09| 0.11| [-0.32, 0.13]   | -0.04   | .423 |
| Points                                    | 3.60 | 0.28| [3.06, 4.14]    | 3.45    | < .001 |
| Point difference                          | 0.07 | 0.13| [-0.18, 0.33]   | 0.23    | .576 |
| **Model Fit**                             |      |     |                 |         |      |
|                                           |      |     | $R^2 = .27$     |         |      |
| **State–situation congruence: Agreeableness and Adversity** |      |     |                 |         |      |
| Intercept                                 | 0.08 | 0.04| [0.01, 0.15]    | 0.28    | .026 |
| State A                                   | 0.23 | 0.06| [0.11, 0.35]    | 0.35    | < .001 |
| Adversity                                 | 0.05 | 0.06| [-0.07, 0.17]   | 0.08    | .420 |
| State $A^2$                               | 0.03 | 0.02| [-0.01, 0.08]   | 0.11    | .108 |
| State A x Adversity                       | 0.02 | 0.03| [-0.04, 0.07]   | 0.06    | .546 |
| Adversity $^2$                            | 0.03 | 0.02| [-0.01, 0.08]   | 0.11    | .187 |
| Behavior conditions                       | 0.20 | 0.11| [-0.02, 0.41]   | 0.09    | .075 |
| Situation conditions                      | -0.04| 0.11| [-0.25, 0.18]   | -0.02   | .743 |
| Points                                    | 2.81 | 0.27| [2.29, 3.34]    | 2.70    | < .001 |
| Point difference                          | 0.28 | 0.10| [0.09, 0.47]    | 0.43    | .004 |
| **Model Fit**                             |      |     |                 |         |      |
|                                           |      |     | $R^2 = .32$     |         |      |
| **State–situation congruence: Agreeableness and Deception** |      |     |                 |         |      |
| Intercept                                 | 0.01 | 0.04| [-0.06, 0.08]   | 0.03    | .749 |
| State A                                   | 0.26 | 0.05| [0.16, 0.36]    | 0.40    | < .001 |
| Deception                                 | 0.07 | 0.06| [-0.05, 0.19]   | 0.10    | .254 |
| State $A^2$                               | 0.04 | 0.02| [-0.01, 0.08]   | 0.12    | .088 |
| State A x Deception                       | -0.02| 0.02| [-0.07, 0.03]   | -0.07   | .372 |
| Deception $^2$                            | 0.00 | 0.02| [-0.05, 0.04]   | -0.01   | .882 |
| Behavior conditions                       | 0.17 | 0.12| [-0.06, 0.40]   | 0.08    | .152 |
| Situation conditions                      | -0.08| 0.11| [-0.30, 0.15]   | -0.04   | .504 |
| Points                                    | 3.22 | 0.29| [2.65, 3.80]    | 3.09    | < .001 |
between personality traits and personality states and congruence between personality states and situation characteristics was not associated with tiredness.

### Associations Between Congruence and Stroop Performance

Effects in the Stroop task are often found either regarding reaction times or regarding error rates (Zelenski et al., 2012). We therefore examined both dependent variables separately using specification-curve analyses (Simonsohn et al., 2019).

#### Stroop Effect

As expected, we found significant Stroop effects in the numerical Stroop task. Participants had significantly longer reaction times in incongruent trials than in congruent trials (median difference across all specifications: 106.71ms) and made more errors in incongruent trials than in congruent trials (median difference across all specifications: 3%). We found these significant Stroop effects in all specifications. Moreover, permutation tests demonstrated that the null hypothesis of no Stroop effect had to be rejected across the whole specification curve, $p < .001$ for reaction times and $p = .014$ for error rates. More details on the Stroop effect can be found in the OSM.

#### Associations Between Congruence and Reaction Times

We tested the congruence hypotheses both with analyses of variance including interactions between median-split personality traits, behavior conditions, and situation conditions and with response surface analyses examining interactions and fit patterns between self-reported personality traits, personality states, and situation characteristics. Here, we summarize the overall pattern of results. Detailed tables and figures concerning these analyses can be found in the OSM.

In a first step, we conducted separate analyses of variance and response surface analyses for all specifications of the dependent variables and examined interaction effects and fit patterns across the different specifications in de-
Figure 3. Response surface plots for associations between trait–state congruence (Panels A and B) and state–situation congruence (Panels C to F) and predicted values of positive affect

Note. Situation characteristics were reverse-coded (indicated by \((r)\)) such that high values represent more desirable values and are congruent with high values of the personality states. The response surface parameters are listed below the titles. The blue lines represent the line of congruence (LOC; i.e., \(X = Y\)) and the line of incongruence (LOIC; i.e., \(X = -Y\)).

scriptive specification curves. Figure 5 shows two examples of such descriptive specification curves for the interactions between median-split trait honesty-humility and the behavior conditions (from an analysis of variance) and between self-reported trait honesty-humility and state honesty-humility (from a polynomial regression). Results of the remaining associations can be found in the OSM. Overall, the descriptive specification curves indicated that none or only very few interactions were significantly associated with reaction times in the Stroop task. The only exception was a significant interaction between median-split trait agreeableness and the behavior conditions, which, however, had the wrong sign and thus indicated slower reaction times in congruent conditions.

Figures 5C and 5D indicate which analytic decisions corresponded to the effects depicted in Figures 5A and 5B. This view can be used to understand how the analytic decisions affect the size of the effect. Only the decision which conditions (incongruent trials only, congruent trials only, neutral trials only, or all trials) were used to calculate reaction times consistently affected the size of the effect: Interaction coefficients were typically lowest when only neutral trials were analyzed and highest when only incongruent trials were analyzed. In contrast, neither the inclusion of results of the game as covariates nor the summary statistics or choice of cutoffs had a systematic effect on the results.

In the second step, the descriptive specification curves were compared to under-the-null specification curves in permutations tests. Table 5 summarizes the results from these permutation tests for all analyses. Apart from the already discussed significant but wrong-directed interaction, none of the permutation tests for reaction times were significant. That is, the result supported the null hypothesis of no effect across the specification curve. Moreover, Fig-
Table 4. Results from polynomial regression models of association between tiredness and congruence

| Term                                      | $b$   | SE  | 95% CI         | $\beta$ | $p$   |
|-------------------------------------------|-------|-----|----------------|---------|-------|
| **Trait–state congruence: Agreeableness** |       |     |                |         |       |
| Intercept                                 | 3.13  | 0.39| [2.37, 3.90]   | 3.42    | < .001|
| State A                                   | 0.22  | 0.04| [0.13, 0.30]   | 0.37    | < .001|
| Trait A                                   | 0.11  | 0.06| [-0.01, 0.23]  | 0.11    | .062  |
| State A$^2$                                | 0.02  | 0.02| [-0.01, 0.06]  | 0.08    | .224  |
| State A x Trait A                         | -0.04 | 0.03| [-0.11, 0.02]  | -0.08   | .190  |
| Trait A$^2$                                | 0.06  | 0.04| [-0.01, 0.13]  | 0.08    | .089  |
| Behavior conditions                       | 0.15  | 0.12| [-0.08, 0.38]  | 0.08    | .193  |
| Situation conditions                      | -0.08 | 0.21| [-0.50, 0.33]  | -0.05   | .694  |
| Points                                    | 0.00  | 0.00| [-0.01, 0.00]  | -0.05   | .827  |
| Point difference                          | 0.00  | 0.00| [0.00, 0.00]   | 0.17    | .369  |
| Model Fit                                 |       |     |                |         | R$^2$ = .20 |
| **Trait–state congruence: Honesty-Humility** |       |     |                |         |       |
| Intercept                                 | 3.44  | 0.38| [2.69, 4.19]   | 3.75    | < .001|
| State HH                                  | 0.12  | 0.06| [0.01, 0.24]   | 0.26    | .039  |
| Trait HH                                  | -0.15 | 0.08| [-0.32, 0.01]  | -0.19   | .073  |
| State HH$^2$                              | -0.01 | 0.02| [-0.04, 0.03]  | -0.03   | .686  |
| State HH x Trait HH                       | 0.05  | 0.03| [-0.01, 0.10]  | 0.21    | .082  |
| Trait HH$^2$                              | 0.07  | 0.03| [0.00, 0.13]   | 0.21    | .043  |
| Behavior conditions                       | 0.11  | 0.12| [-0.13, 0.34]  | 0.06    | .380  |
| Situation conditions                      | -0.26 | 0.19| [-0.62, 0.11]  | -0.14   | .171  |
| Points                                    | 0.00  | 0.00| [0.00, 0.00]   | 0.04    | .858  |
| Point difference                          | 0.00  | 0.00| [0.00, 0.00]   | 0.18    | .303  |
| Model Fit                                 |       |     |                |         | R$^2$ = .18 |
| **State–situation congruence: Agreeableness and Adversity** |       |     |                |         |       |
| Intercept                                 | 3.15  | 0.40| [2.37, 3.94]   | 3.44    | < .001|
| State A                                   | 0.16  | 0.05| [0.05, 0.26]   | 0.27    | .004  |
| Adversity                                 | 0.00  | 0.06| [-0.11, 0.11]  | 0.00    | .979  |
| State A$^2$                                | 0.02  | 0.02| [-0.02, 0.06]  | 0.05    | .443  |
| State A x adversity                       | 0.03  | 0.03| [-0.02, 0.09]  | 0.14    | .203  |
| Adversity$^2$                             | 0.03  | 0.02| [-0.02, 0.07]  | 0.10    | .228  |
| Behavior conditions                       | 0.15  | 0.11| [-0.07, 0.38]  | 0.08    | .170  |
| Situation conditions                      | -0.13 | 0.20| [-0.52, 0.27]  | -0.07   | .532  |
| Points                                    | 0.00  | 0.00| [-0.01, 0.00]  | -0.06   | .772  |
| Point difference                          | 0.00  | 0.00| [0.00, 0.00]   | 0.16    | .384  |
| Model Fit                                 |       |     |                |         | R$^2$ = .21 |
| **State–situation congruence: Agreeableness and Deception** |       |     |                |         |       |
| Intercept                                 | 3.14  | 0.40| [2.36, 3.93]   | 3.43    | < .001|
| State A                                   | 0.24  | 0.05| [0.15, 0.34]   | 0.42    | < .001|
| Deception                                 | 0.08  | 0.06| [-0.03, 0.19]  | 0.13    | .150  |
| State A$^2$                                | 0.01  | 0.02| [-0.03, 0.05]  | 0.05    | .484  |
| State A x Deception                       | 0.01  | 0.02| [-0.04, 0.06]  | 0.03    | .758  |
| Deception$^2$                             | 0.03  | 0.02| [-0.01, 0.08]  | 0.13    | .107  |
| Behavior conditions                       | 0.14  | 0.12| [-0.09, 0.37]  | 0.08    | .237  |
| Situation conditions                      | -0.12 | 0.21| [-0.53, 0.29]  | -0.06   | .566  |
| Points                                    | 0.00  | 0.00| [-0.01, 0.00]  | -0.01   | .949  |
Thus, neither trait–state congruence nor state–situation congruence was significantly associated with error rates in the Stroop task. Moreover, none of the analytic decisions consistently affected the size or direction of the interaction effects (Figures 6C and 6D and additional figures in the OSM).

Permutation tests confirmed that neither interactions nor fit patterns between personality traits and behavior conditions, between situation and behavior conditions, between personality traits and personality states, or between personality states and situation characteristics were associated with error rates in the Stroop task (Table 5). P values of the permutation test were all bigger than \( p = .030 \) and the effects in the real data were well within the 2.5th and 97.5th percentiles of the effects found in the shuffled datasets (Figures 6E and 6F). Thus, neither trait–state congruence nor state–situation congruence was significantly associated with error rates in the Stroop task.

**Discussion**

We examined whether two forms of behavioral congruence, trait–state congruence and state–situation congruence, were associated with positive affect, tiredness, and performance in a Stroop task. Overall, the results led to three main findings: First, the manipulation check demonstrated that although the manipulations had the expected effects, tiredness was reverse-coded such that higher values indicate less tiredness or a more active mood. Personality traits, personality states, and situation characteristics were centered on the midpoints of their respective scales. Personality traits were transformed from a 5-point scale to a 7-point scale.

### Associations Between Congruence and Error Rates

We repeated the specification curve analyses with error rates as the dependent variables. Figures 6A and 6B again show two examples of descriptive specification curves, the remaining figures can be found in the OSM. Overall, none or only very few interaction coefficients were significantly associated with error rates in the Stroop task. Moreover, none of the analytic decisions consistently affected the size or direction of the interaction effects (Figures 6C and 6D and additional figures in the OSM).

| Term                        | \( b \) | SE  | 95% CI            | \( \beta \) | \( p \)  |
|-----------------------------|--------|-----|-------------------|-------------|--------|
| Point difference            | 0.00   | 0.00| [0.00, 0.00]      | 0.12        | .523   |
| Model Fit                   |        |     | \( R^2 = .20 \)   |             |        |

State–situation congruence: Honesty-Humility and Adversity

| Term                        | \( b \) | SE  | 95% CI            | \( \beta \) | \( p \)  |
|-----------------------------|--------|-----|-------------------|-------------|--------|
| Intercept                   | 3.49   | 0.39| [2.72, 4.26]      | 3.81        | < .001 |
| State HH                    | 0.14   | 0.05| [0.04, 0.25]      | 0.30        | .008   |
| Adversity                   | 0.02   | 0.06| [-0.10, 0.13]     | 0.03        | .797   |
| State HH\(^2\)              | -0.02  | 0.02| [-0.05, 0.02]     | -0.06       | .362   |
| State HH x Adversity        | 0.03   | 0.02| [-0.01, 0.07]     | 0.16        | .150   |
| Adversity\(^2\)             | 0.04   | 0.02| [0.00, 0.09]      | 0.17        | .076   |
| Behavior conditions         | 0.06   | 0.12| [-0.17, 0.29]     | 0.03        | .625   |
| Situation conditions        | -0.16  | 0.19| [-0.53, 0.21]     | -0.09       | .400   |
| Points                      | 0.00   | 0.00| [-0.01, 0.00]     | -0.08       | .699   |
| Point difference            | 0.00   | 0.00| [0.00, 0.00]      | 0.22        | .234   |
| Model Fit                   |        |     | \( R^2 = .20 \)   |             |        |

State–situation congruence: Honesty-Humility and Deception

| Term                        | \( b \) | SE  | 95% CI            | \( \beta \) | \( p \)  |
|-----------------------------|--------|-----|-------------------|-------------|--------|
| Intercept                   | 3.44   | 0.38| [2.70, 4.18]      | 3.74        | < .001 |
| State HH                    | 0.25   | 0.05| [0.16, 0.35]      | 0.53        | < .001 |
| Deception                   | 0.08   | 0.05| [-0.02, 0.17]     | 0.12        | .105   |
| State HH\(^2\)              | -0.01  | 0.02| [-0.05, 0.02]     | -0.04       | .502   |
| State HH x Deception        | 0.02   | 0.02| [-0.02, 0.07]     | 0.09        | .343   |
| Deception\(^2\)             | 0.05   | 0.02| [0.01, 0.09]      | 0.19        | .012   |
| Behavior conditions         | 0.03   | 0.12| [-0.21, 0.27]     | 0.01        | .828   |
| Situation conditions        | -0.14  | 0.19| [-0.51, 0.23]     | -0.07       | .468   |
| Points                      | 0.00   | 0.00| [0.00, 0.00]      | -0.01       | .948   |
| Point difference            | 0.00   | 0.00| [0.00, 0.00]      | 0.16        | .366   |
| Model Fit                   |        |     | \( R^2 = .18 \)   |             |        |

Note. The lines separate the different models from each other. Tiredness was reverse-coded such that higher values indicate less tiredness or a more active mood. Personality traits, personality states, and situation characteristics were centered on the midpoints of their respective scales. Personality traits were transformed from a 5-point scale to a 7-point scale. Trait A and State A refer to trait and state agreeableness, respectively, and Trait HH and State HH refer to trait and state honesty-humility, respectively.
effects, they also had unintended unspecific effects. Second, both analyses of variance with experimental conditions and response surface analyses with self-reported situation characteristics and personality states did not yield any trait–state congruence or state–situation congruence effects. Third, positive affect and tiredness were only consistently associated with behavior (i.e., experimental behavior conditions or personality states), whereas Stroop performance was not associated with any of the predictors.

**Experimental Manipulations**

This study was intended to experimentally manipulate trait–state congruence and state–situation congruence to examine their effects on mood and cognitive performance. We aimed to manipulate the behavior and personality states in a prisoner’s dilemma game by directly instructing participants to act either honestly and agreeably or dishonestly and disagreeably. Direct instructions as manipulations of personality states have already been used in a few studies (e.g., McNiel et al., 2010; Zelenksi et al., 2012) and we therefore expected them to be effective. Interestingly, the instructions did have the intended effects on state agreeableness and state honesty-humility, but they also had unintended effects on state openness, state conscientiousness, and state emotionality. The manipulation had therefore rather unspecific effects, affecting not only the targeted personality states but also most of the other personality states. As far as we are aware, previous studies using direct instructions to manipulate personality states did not report whether their manipulations also affected other personality states and it is therefore unclear whether this is a general problem with this type of manipulation or specific to our study. If such manipulations indeed affected not
only the desired variables but also undesired variables, this would be a serious threat to the internal validity of these designs. We therefore strongly recommend that future studies using similar manipulations examine their effects on all personality states to test the specificity of the manipulation.

Additionally, we aimed to manipulate perceived adversity and deception of the game by manipulating the game’s payoff structure and the description and behavior of the computer-simulated player. To our best knowledge, no one has manipulated such economic games with the intent to manipulate perceived characteristics of the game before. Our manipulation was therefore novel and exploratory. We observed two main problems with the situation manipulation: First, although the manipulation was significantly associated with the perception of adversity (but not of deception), the effect was rather small. The effect of the manipulation therefore might have been too small to show the hypothesized effects even if the other problems had not existed. Second, as situation perceptions and personality states were interrelated, the effects of the manipulations often showed trends towards interactions between the behavior and situation conditions. Thus, it may be particularly difficult to manipulate situation perception and personality states independently from each other.

Figure 5. Results from specification curve analyses of the associations between the interaction of trait honesty-humility and state honesty-humility (A, C, E) and of state honesty-humility and reverse-coded adversity (B, D, F) and reaction times in the Stroop task

Note. A and B represent descriptive specification curves with specifications ranked by size and significant specifications marked in blue. C and D present the analytics choices that led to the effects displayed in A and B. E and F compare the effect observed in the real data (blue and grey dots) to the 2.5th and 97.5th percentiles (dashed lines) of the effects observed in the shuffled data.
Table 5. Summary of the joint significance tests from specification curve analyses

| Analysis        | Operationalization          | Relevant predictor                               | Number of specifications | Median effect size | Significant specifications (%) | Number of shuffled samples with more significant specifications than for the original sample | p value of permutation test |
|-----------------|-----------------------------|-------------------------------------------------|--------------------------|-------------------|-------------------------------|------------------------------------------------------------------------------------------|---------------------------|
| **DV: Reaction times** |                             |                                                 |                          |                   |                               |                                                                                          |                           |
| AOV Interaction | Trait A x Behavior Conditions | 8,896                                           | 114                      | 66%               | 9                             | .018                                                                                      |                           |
| AOV Interaction | Trait H x Behavior Conditions | 8,896                                           | -11.4                    | 0%                | 500                           | > .999                                                                                     |                           |
| AOV Interaction | Situation x Behavior Conditions | 8,896                                           | 48.61                    | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Interaction | Trait A x State A            | 8,896                                           | -1.15                    | 0%                | 400                           | > .999                                                                                     |                           |
| RSA Interaction | Trait HH x State HH          | 8,896                                           | -5.02                    | 3%                | 105                           | .210                                                                                      |                           |
| RSA Interaction | State A x Adversity          | 8,896                                           | 3.07                     | 0%                | 192                           | > .999                                                                                     |                           |
| RSA Interaction | State A x Deception          | 8,896                                           | -4.25                    | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Interaction | State HH x Adversity         | 8,896                                           | -2.36                    | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Interaction | State HH x Deception         | 8,896                                           | -0.58                    | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Fit pattern | Trait A x State A            | 8,896                                           | no                       | 0%                | 161                           | .402                                                                                      |                           |
| RSA Fit pattern | Trait HH x State HH          | 8,896                                           | no                       | 1%                | 89                            | .178                                                                                      |                           |
| RSA Fit pattern | State A x Adversity          | 8,896                                           | no                       | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Fit pattern | State A x Deception          | 8,896                                           | no                       | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Fit pattern | State HH x Adversity         | 8,896                                           | no                       | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Fit pattern | State HH x Deception         | 8,896                                           | no                       | 0%                | 500                           | > .999                                                                                     |                           |
| **DV: Error rates** |                             |                                                 |                          |                   |                               |                                                                                          |                           |
| AOV Interaction | Trait A x Behavior Conditions | 2,224                                           | -0.01                    | 5%                | 127                           | .254                                                                                      |                           |
| AOV Interaction | Trait H x Behavior Conditions | 2,224                                           | 0                       | 0%                | 376                           | .752                                                                                      |                           |
| AOV Interaction | Situation x Behavior Conditions | 2,224                                           | 0                       | 0%                | 370                           | .740                                                                                      |                           |
| RSA Interaction | Trait A x State A            | 2,224                                           | 0                       | 1%                | 274                           | .548                                                                                      |                           |
| RSA Interaction | Trait HH x State HH          | 2,224                                           | 0                       | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Interaction | State A x Adversity          | 2,224                                           | 0                       | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Interaction | State A x Deception          | 2,224                                           | 0                       | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Interaction | State HH x Adversity         | 2,224                                           | 0                       | 0%                | 441                           | .882                                                                                      |                           |
| RSA Interaction | State HH x Deception         | 2,224                                           | 0                       | 0%                | 396                           | .792                                                                                      |                           |
| RSA Fit pattern | Trait A x State A            | 2,224                                           | no                      | 2%                | 217                           | .434                                                                                      |                           |
| RSA Fit pattern | Trait HH x State HH          | 2,224                                           | no                      | 7%                | 89                            | .178                                                                                      |                           |
| RSA Fit pattern | State A x Adversity          | 2,224                                           | no                      | 0%                | 500                           | > .999                                                                                     |                           |
| RSA Fit pattern | State A x Deception          | 2,224                                           | no                      | 20%               | 30                            | .060                                                                                      |                           |
| RSA Fit pattern | State HH x Adversity         | 2,224                                           | no                      | 0%                | 416                           | .832                                                                                      |                           |
| RSA Fit pattern | State HH x Deception         | 2,224                                           | no                      | 25%               | 15                            | .030                                                                                      |                           |

Note. Trait A and State A refer to trait and state agreeableness, respectively, and Trait HH and State HH refer to trait and state honesty-humility, respectively. AOV represent analyses of variance, RSA represent response surface analyses. For fit patterns, the median effect size is indicated as presence of a fit pattern (yes) or absence of a fit pattern (no).
Moreover, the ulterior motive of these manipulations was to manipulate congruence between personality states and personality trait and between personality states and situation characteristics. Descriptive analyses (see OSM) showed that difference scores between the variables (used as an approximation of congruence) differed less between the experimental conditions than the manipulated variables themselves. For example, although state agreeableness was significantly lower in the low agreeableness and honesty condition, participants on average behaved more agreeably than their trait level in all experimental conditions. However, these difference scores should be interpreted with caution because they assume strong equivalence between the trait and state scales (e.g., Schönbrodt, 2016) which may not be given in this study. Overall, experimentally manipulating congruence may thus be even more challenging as it may require more than just manipulating the involved variables (J. R. Edwards et al., 2006) and may be more difficult to prove.

Despite the many difficulties in manipulating personality states, situation characteristics, and trait–state and state–situation congruence we have already discussed, we are convinced that such manipulations are important. Experimental studies with valid and reliable manipulations can further the understanding of personality in many ways. In the context of this study, it is particularly important to note that such experimental studies could deepen our understanding of personality processes and the causal relationships between personality states and other constructs such as well-being, motivation, or authenticity (e.g., Fleeson & Law, 2015; McCabe & Fleeson, 2016). Furthermore, such research could also help to inform the validity of manipulations in the existing literature (e.g., Fleeson & Wilt, 2010; Jacques-Hamilton et al., 2018; Margolis & Lyubomirsky, 2020; McCabe & Fleeson, 2016). Finally, experimental studies that manipulate personality states and situation characteristics independently from each other may provide important insight into their associations and
interactions.

**Associations Between Congruence and Positive Affect and Tiredness**

We examined how interactions and fit patterns as operationalizations of congruence were associated with positive affect and tiredness. Theories such as trait-activation theory (Tett & Burnett, 2003; Tett & Guterman, 2000) and the contra-trait effort hypothesis (Gallagher et al., 2011) predicted that behaving congruently should be associated with more positive affect and less tiredness, whereas behaving incongruently should be associated with less positive affect and more tiredness. However, we found that neither interactions nor fit patterns were associated with either positive affect or tiredness. Across all models, only state agreeableness and state honesty-humility were consistently significantly associated with positive affect and tiredness.

These findings have both theoretical and practical implications: First, our findings are consistent with those from previous studies that also did not find any interactions or congruence effects but only main effects of personality states (e.g., Fleeson et al., 2002; Fleeson & Wilt, 2010; McNiel et al., 2010; Zelenski et al., 2012). This supports the concept of trait-state isomorphism (Fleeson et al., 2002): Personality states are functionally equivalent to personality traits, that is, they have the same correlates and consequences. Similarly, research on congruence and authenticity has also pointed towards a state-content significance hypothesis (Fleeson & Wilt, 2010) such that only the absolute levels of personality states were associated with authenticity, but not their congruence with personality traits. Future research should thus further examine these ideas of trait-state isomorphism (e.g., van Allen & Zelenski, 2018) and state-content significance.

However, it is also possible that our sample might have been too small to detect congruence effects. Within-situation effects are typically rather small and unfold their power by accumulating over many situations (Funder & Ozer, 2019; Rauthmann et al., 2015). Particularly interactions and fit patterns are statistically more complex and thus require a bigger sample size than main effects to be reliably detected. We used a previously reported effect size for power considerations for this study (Zelenski et al., 2012), but this effect was based on an interaction effect between behavioral conditions and extreme groups of trait extraversion in an analysis of variance. It may therefore not have been the best benchmark for the analyses we conducted, particularly for the response surface analyses.

Notably, significant congruence effects were most often reported in experience sampling studies (e.g., Jacques-Hamilton et al., 2018; Pickett, Hofmans, Debusscher, et al., 2020; Pickett, Hofmans, Feldt, et al., 2020; but see also Margolis & Lyubomirsky, 2020), whereas null effects were most often reported in experimental lab studies (e.g., Fleeson et al., 2002; McNiel et al., 2010; McNiel & Fleeson, 2006). One of the systematic differences between these types of studies were their sample sizes with experimental studies typically examining at most 200 participants and experience sampling studies typically examining more than 1,000 different observations. Thus, one might suspect that these differences in sample sizes (and accordingly in statistical power) might at least partially explain the differences in findings.

**Associations Between Congruence and Stroop Performance**

We also examined how interactions and fit patterns as operationalizations of congruence were associated with reaction times and error rates in a Stroop task. Theories such as the contra-trait effort hypothesis (Gallagher et al., 2011) suggested that behaving incongruently depletes self-control and that this self-control is then unavailable for subsequent tasks requiring mental effort. However, we found that neither interactions nor fit patterns were associated with reaction times or error rates in the Stroop task.

These results can be interpreted in different ways: First, the hypothesis may simply be false and subsequent performance is not impaired by incongruent personality states. As this hypothesis is based on the ego depletion literature (Baumeister et al., 1998, 2007) which has recently been heavily criticized (Carter et al., 2015; Friese et al., 2019; Hagger et al., 2016; Inzlicht & Friese, 2019; Luquini et al., 2016), there may not be a sound theoretical basis for hypothesizing that congruence is associated with cognitive performance. Second, self-control includes different responses such as controlling thoughts, regulating emotions, or guiding behavior (Baumeister et al., 2007) and the task may not have measured the aspects of self-control that are impacted by incongruent personality states (Gallagher et al., 2011). Third, it could be argued that a certain amount or duration of incongruent personality states is needed to exert enough self-control to have an effect (Gallagher et al., 2011). Our situation may have been too short or the incongruence too small to see the hypothesized effect.

Overall, we did not find any evidence that trait-state congruence or state-situation congruence were significantly associated with cognitive performance in a Stroop task. There are several possible explanations, and based on our study, we unfortunately cannot conclude which of these explanations is correct. However, since the theoretical basis for these hypotheses is currently under heavy criticism, future research seems best advised to wait and see what the ego depletion debate yields before further applying these theories to incongruent personality states.

**Importance of Statistical Operationalizations**

Our study emphasizes the importance of operationalizations and analytic decisions. Congruence is a concept that has been examined using a multitude of operationalizations from difference scores to interactions to polynomial regressions and response surface analyses (e.g., J. R. Edwards, 1993; Humberg, Dufner, et al., 2019). In our study, we found that although the grand conclusions were the same regardless of the operationalizations, operationalizing congruence as interaction effects and fit patterns did not always yield the exact same results. As discussed before, a significant interaction coefficient, for example, only indicates that the relationship between two variables depends on the level of a third variable. This alone does not necessarily indicate a congruence effect. Therefore, it is important to carefully...
decide on the operationalization to use and to be aware of the implications of these decisions. Additionally, the theoretical concept of congruence needs to be further refined so that it becomes clearer which operationalizations are appropriate under which circumstances and which empirical findings are informative for these theories, and which are not. As response surface analyses include and differentiate between the different operationalizations (e.g., Schönbrodt, 2016), it could be a valuable instrument in this endeavor.

Moreover, our study also emphasized the importance of analytic decisions and being aware of researcher degrees of freedom (Simmons et al., 2011). We conducted specification curve analyses (Simonsohn et al., 2020) to account for the multitude of possible analytic decisions regarding the calculation of reaction time and error rate indices in the Stroop task. Specification curve analyses can reveal whether an effect (a) is present across the majority of specifications and thus quite robust, (b) is only present in a subset of specifications and thus depends on known or unknown boundary conditions, or (c) is not present across the majority of specifications and thus most likely does not exist. In our specification curve analyses, the great majority of specifications did not yield any significant congruence or interaction effects and we thus concluded that there is most likely no association between behavioral congruence and performance in the Stroop task. However, we still found that the size of the effects depended on the analytic decisions. Particularly the decision over which type of trials the reaction times were computed (i.e., all trials vs. only congruent/incongruent/neutral trials) was associated with the size of the interaction effects. Additionally, there were some specifications, albeit very few, that resulted in significant effects. These were not enough to lead to a different conclusion, but if we had arbitrarily selected just one specification for the analyses and incidentally selected one of these specifications with a significant effect, our conclusions would have been different. Therefore, it is important to test the complete specification curve to be able to conclude that these are most likely just false positives. Overall, then, this study emphasizes that researchers should be aware of the importance and implications of their decisions and actively incorporate the implications of these decisions into their research more often—for example by using specification curve analyses (Simonsohn et al., 2020).

Limitations

Beyond the well-known and often discussed limitations of self-reports and WEIRD samples (Henrich et al., 2010), this study had a few rather specific limitations. First, the study was conducted online, which allowed us to collect a bigger and more diverse sample than possible in the lab, but also led to complications. We had, for example, less control over the participants and their environments and possible distractions in such an online study. This might be particularly important when cognitive tasks such as the Stroop task are included. Moreover, situation research in an online study may also bring about some challenges because it is unknown how comparable online situations are to "real" situations in the lab or even in everyday life. It is therefore possible that participants would have perceived and reacted to a comparable prisoner’s dilemma game in the lab differently than to the online game. Researching situations online could be a great opportunity for situation research as it makes larger and more diverse samples accessible, however, future research should aim to learn more about differences and similarities of situations and situation perception online and offline.

Second, we did not assess whether participants believed to be playing the prisoner’s dilemma game against a real person and therefore could not control whether this belief (or differences in the beliefs among the participants) affected their behavior during the game. Whereas there is evidence that economic games work similar on the internet as in the lab (Amir et al., 2012; Baumert et al., 2014) and even when playing against virtual humans or real humans (Nouri & Traum, 2013), there is also some evidence that beliefs about the other players affects behavior in economic games (Eckel & Wilson, 2006; Johnson & Mislin, 2011). Third, as discussed above, we determined the sample size with a power analysis based on a previously reported effect size (Zelenski et al., 2012) but this effect size may not have been the best benchmark for our analyses. Particularly the response surface analyses may have suffered from a loss of power to detect higher order effects such as interactions or quadratic effects. Finally, although internal consistencies are expected to be lower in short scales that aim to increase coverage of the construct, the alpha coefficients for the personality traits Emotionality and Agreeableness were unusually low.

Conclusion

The present study aimed to examine effects of trait–state and state–situation congruence on mood and cognitive performance. The resulting conclusions and future directions are twofold: First, we found that neither trait–state congruence nor state–situation congruence were associated with positive affect, tiredness, or performance in a Stroop task. Instead, positive affect and tiredness were only associated with personality states. Stroop performance was not consistently associated with any of the predictors at all. This study thus seems to support hypotheses like trait–state isomorphism or state content significance rather than congruence hypotheses. However, in light of the limitations of this study, more studies that are carefully designed, carefully operationalized, and well-powered are needed to examine effects of congruence on relevant outcomes such as mood and tiredness—particularly in personality domains other than extraversion. Second, as the manipulations had rather unspecific effects on personality states and perceived situation characteristics, our study emphasized the difficulty of manipulating personality states, situation characteristics, and trait–state and state–situation congruence in general. We need experimental research on personality states and situation characteristics to advance the understanding of personality processes, situations, and their interplay, and to do that we need good manipulations. Thus, although it will likely be a challenging endeavor, future research should aim to establish valid and reliable manipulations of personality states, perceived situation characteris-
tics, and trait–state and state–situation congruence.

Contributions

Contributed to conception and design: SK, ML
Contributed to acquisition of data: SK, ML
Contributed to analysis and interpretation of data: SK
Drafted and/or revised the article: SK, ML
Approved the submitted version for publication: SK, ML

Acknowledgements

We thank our research assistants Charlotte Michel, Anke Kirchdörfer, and Marie von Rögal for their help with the literature research and data collection. Further, we thank Peter Haehner for useful comments on an earlier version of this manuscript. Finally, we acknowledge support by the Open Access Publication Funds of the Ruhr University Bochum.

Competing Interests

The authors have no competing interests to declare.

Supplemental Material

Online Supplementary Material (OSM). HTML document.
Deviations from the preregistration. PDF document.

Data Accessibility Statement

We made all our materials, including data, analysis code, and the reproducible manuscript, publicly available at the Open Science Framework (Link: https://osf.io/8tbep/). The preregistration for this study is also available at the Open Science Framework (Link: https://osf.io/hnu4b/).

Submitted: March 25, 2021 PDT, Accepted: August 04, 2021 PDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at http://creativecommons.org/licenses/by/4.0 and legal code at http://creativecommons.org/licenses/by/4.0/legalcode for more information.
REFERENCES

Amir, O., Rand, D. G., & Gal, Y. K. (2012). Economic games on the internet: The effect of $1 stakes. *PloS One, 7*(2), e31461. https://doi.org/10.1371/journal.pone.0031461

Ashton, M. C., & Lee, K. (2001). A theoretical basis for the major dimensions of personality. *European Journal of Personality, 15*(5), 327–355. https://doi.org/10.1002/per.417

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., De Vries, R. E., & de Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholinguistic studies in seven languages. *Journal of Personality and Social Psychology, 86*(2), 356–366. https://doi.org/10.1037/0022-3514.86.2.356

Aust, F., & Barth, M. (2020). *papaja: Prepare reproducible APA journal articles with R Markdown*. https://github.com/crsh/papaja

Barth, M. (2021). *tinylabels: Lightweight variable labels*. https://github.com/mariusbarth/tinylabels

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., & Maechler, M. (2019). *Matrix: Sparse and dense matrix classes and methods*. https://CRAN.R-project.org/package=Matrix

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74*(5), 1252–1265. https://doi.org/10.1037/0022-3514.74.5.1252

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The HEXACO honesty-humility, agreeableness, and emotionality factors: A review of research and theory. *Journal of Personality, 75*(5), 327–353. https://doi.org/10.1111/j.1467-8826.2006.00534.x

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General, 144*(4), 796–815. https://doi.org/10.1037/0097-7403.144.4.796

Chandler, J., Rosenzweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online panels in social science research: Expanding sampling methods beyond mechanical turk. *Behavior Research Methods, 51*(5), 2022–2038. https://doi.org/10.3758/s13428-019-01273-7

Ching, C. M., Church, A. T., Katigbak, M. S., Reyes, J. A. S., Tanaka-Matsumi, J., Takooka, S., Zhang, H., Shen, J., Arias, R. M., Rincon, B. C., & Ortiz, F. A. (2014). The manifestation of traits in everyday behavior and affect: A five-culture study. *Journal of Research in Personality, 48*, 1–16. https://doi.org/10.1016/j.jrp.2013.10.002

Churchyard, J. S. (2013). *Within-person variation in personality and psychological well-being* [Dissertation, University of Hertfordshire]. https://doi.org/10.18745/th.15452

Cooper, A. B., Sherman, R. A., Rauthmann, J. F., Serfass, D. G., & Brown, N. A. (2018). Feeling good and authentic: Experienced authenticity in daily life is predicted by positive feelings and situation characteristics, not trait-state consistency. *Journal of Research in Personality, 77*, 57–69. https://doi.org/10.1016/j.jrp.2018.09.005

Cronbach, L. J. (1958). Proposals leading to analytic treatment of social perception scores. *Person Perception and Interpersonal Behavior, 53*(579).

De Vries, R. E. (2013). The 24-item brief HEXACO inventory (BHI). *Journal of Research in Personality, 47*(6), 871–880. https://doi.org/10.1016/j.jrp.2013.09.003

Diener, E., & Lucas, R. E. (2020). Personality traits. In R. Biswas-Diener & E. Diener (Eds.), *Noba textbook series: psychology*. DEF Publishers.

Eckel, C. C., & Wilson, R. K. (2006). Internet cautions: Experimental games with internet partners. *Experimental Economics, 9*(1), 53–66. https://doi.org/10.1007/s10683-006-4307-4

Edwards, J. R. (1993). Problems with the use of profile similarity indices in the study of congruence in organizational research. *Personnel Psychology, 46*(3), 641–665. https://doi.org/10.1177/10889789.1993.tb00889.x

Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Vols. Advances in Measurement and Data Analysis* (pp. 350–400). Jossey-Bass.

Edwards, J. R. (2007). Polynomial regression and response surface methodology. *Perspectives on Organizational Fit, 561–372.

Edwards, J. R., Cable, D. M., Williamson, I. O., Lambert, L. S., & Shipp, A. J. (2006). The phenomenology of fit: Linking the person and environment to the subjective experience of person-environment fit. *Journal of Applied Psychology, 91*(4), 802–827. https://doi.org/10.1037/0021-9010.91.4.802
Funder, D. C. (2016). Taking situations seriously. Social Psychological and Personality Science, 24(6), 335–342. https://doi.org/10.1177/1948550610390701

Gamer, M., Lemon, J., & Singh, I. F. P. (2019). Irr: Various coefficients of interrater reliability and agreement. https://CRAN.R-project.org/package=irr

Gu, Z., Gu, L., Eils, R., Schlesner, M., & Brors, B. (2014). Circize implements and enhances circular visualization in r. Bioinformatics, 30(19), 2811–2812. https://doi.org/10.1093/bioinformatics/btu393

Hagjer, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., Brand, R., Brandt, M. J., Brewer, G., BrunyeeL, S., Calvillo, D. P., Campbell, W. K., Cannon, P. R., Carlucci, M., Carruth, N. P., Cheung, T., Crowell, A., de Ridder, T. D. T., Dewitte, S., ... Zwiennenberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. Perspectives on Psychological Science, 11(4), 546–573. https://doi.org/10.1177/1745691616652873

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. Memory & Cognition, 10(4), 389–395. https://doi.org/10.3758/bf03202431

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Beyond WEIRD: Towards a broad-based behavioral science. Behavioral and Brain Sciences, 33(2–5), 111–135. https://doi.org/10.1017/s0140525x1000725

Henry, L., & Wickham, H. (2020). Purrr: Functional programming tools. https://CRAN.R-project.org/package=purrr

Humberg, S., Dufner, M., Schönbrot, F. D., Geukes, K., Hutteman, R., Köfler, A. C. P., van Zalk, M. H. W., Denissen, J. J. A., Nestler, S., & Back, M. D. (2019). Is accurate, positive, or inflated self-perception most advantageous for psychological adjustment? A competitive test of key hypotheses. Journal of Personality and Social Psychology, 116(5), 835–859. https://doi.org/10.1037/pspp0000304

Humberg, S., Nestler, S., & Back, M. D. (2019). Response surface analysis in personality and social psychology: Checklist and clarifications for the case of congruence hypotheses. Social Psychological and Personality Science, 10(5), 409–419. https://doi.org/10.1177/1948550618757600

Inzlicht, M., & Friese, M. (2019). The past, present, and future of ego depletion. Social Psychology, 50(5–6), 370–378. https://doi.org/10.1027/1864-9335/a000398

Jacques-Hamilton, R., Sun, J., & Smillie, L. D. (2018). Costs and benefits of acting extravedted: A randomized controlled trial. Journal of Experimental Psychology: General, 147(9), 1558–1556. https://doi.org/10.1037/xe0000516

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. Journal of Economic Psychology, 32(5), 865–889. https://doi.org/10.1016/j.joep.2011.05.007

Kagel, J., & McGee, P. (2014). Personality and cooperation in finitely repeated prisoner's dilemma games. Economics Letters, 124(2), 274–277. https://doi.org/10.1016/j.econlet.2014.05.034
Rauthmann, J. F., & Sherman, R. A. (2016b). Ultra-brief measures for the situational eight DIAMONDS domains. European Journal of Psychological Assessment, 32(2), 165–174. https://doi.org/10.1027/1015-5759/a000245

Rauthmann, J. F., Sherman, R. A., Nave, C. S., & Funder, D. C. (2015). Personality-driven situation experience, contact, and construal: How people’s personality traits predict characteristics of their situations in daily life. Journal of Research in Personality, 55, 98–111. https://doi.org/10.1016/j.jrp.2015.02.003

Revelle, W. (2021). Psych: Procedures for psychological, psychometric, and personality research. Northwestern University. https://CRAN.R-project.org/package=psych

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. Journal of Statistical Software, 48(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Sarkar, D. (2008). Lattice: Multivariate data visualization with R. Springer. http://lmdvr.r-forge.r-project.org

Schönbrodt, F. D. (2016). Testing fit patterns with polynomial regression models. PsyArXiv. https://doi.org/10.31219/osf.io/ndggf

Schönbrodt, F. D., & Humberg, S. (2021). RSA: An R package for response surface analysis (version 0.10.4). https://cran.r-project.org/package=RSA

Sherman, R. A., Rauthmann, J. F., Nave, C. S., & Funder, D. C. (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. Journal of Personality and Social Psychology, 109(5), 872–888. https://doi.org/10.1037.jspsp0000056

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). A false-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological Science, 22(11), 1539–1536. https://doi.org/10.1177/0956797611417632

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2160588

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2019). Specification curve analysis. Nature Human Behaviour, 4(11), 1208–1214. https://doi.org/10.1038/s41562-020-0912-z

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2021). Afex: Analysis of factorial experiments. https://CRAN.R-project.org/package=afex

Stanley, D. (2021). apaTables: Create american psychological association (APA) style tables. https://CRAN.R-project.org/package=apaTables

Steyer, R. (2021, June 22). MDMQ questionnaire (english version of MDBF). https://www.metheval.uni-jena.de/mdbf.php

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1994). Testtheoretische analysen des mehrdimensionalen behindlichkeitsfragebogen (MDBF) [theoretical analysis of a multidimensional mood questionnaire (MDBF)]. Diagnostica, 40(4), 320–328.

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. Journal of Applied Psychology, 88(3), 500–517. http://dx.doi.org/10.1037/0021-9010.88.3.500

Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. Journal of Research in Personality, 34(4), 397–423. http://dx.doi.org/10.1006/jrpe.2000.2292

Tucker, A. W. (1980). A two-person dilemma. UMAP Journal, 1, 101–105.

Tucker, A. W. (1985). The mathematics of tucker: A sampler. The Two-Year College Mathematics Journal, 14(5), 228. https://doi.org/10.2307/3027092

Urbanke, S. (2019). Jpeg: Read and write JPEG images. https://CRAN.R-project.org/package=jpeg

van Allen, Z. M., & Zelenski, J. M. (2018). Testing trait-state isomorphism in a new domain: An exploratory manipulation of openness to experience. Frontiers in Psychology, 9, 1964. https://doi.org/10.3389/fpsyg.2018.01964

Whelan, D. C. (2015). Extraversion and counter-dispositional behavior: Exploring consequences and the impact of situation-behavior congruence [Dissertation]. Carleton University.

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H. (2019). Stringr: Simple, consistent wrappers for common string operations. https://CRAN.R-project.org/package=stringr

Wickham, H. (2021a). Forcats: Tools for working with categorical variables (factors). https://CRAN.R-project.org/package=forcats

Wickham, H. (2021b). Tidy: Tidy messy data. https://CRAN.R-project.org/package=tidy

Wickham, H., Averbick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grelumond, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(45), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2019). Readxl: Read excel files. https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., & Müller, K. (2021). Dplyr: A grammar of data manipulation. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Hester, J. (2020). Readr: Read rectangular text data. https://CRAN.R-project.org/package=readr

Wilhelm, P., & Schoebi, D. (2007). Assessing mood in daily life. European Journal of Psychological Assessment, 23(4), 258–267. https://doi.org/10.1027/1015-5759.23.4.258
Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'*: [https://CRAN.R-project.org/package=cowplot](https://CRAN.R-project.org/package=cowplot)

Wilt, J. A., Sun, J., Jacques-Hamilton, R., & Smillie, L. D. (2021). *Why does it feel authentic to be and act extraverted? Exploring the mediating role of positive affect*: [https://doi.org/10.31234/osf.io/7mj6g](https://doi.org/10.31234/osf.io/7mj6g)

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Chapman; Hall/CRC. [https://yihui.org/knitr/](https://yihui.org/knitr/)

Zelenski, J. M., Santoro, M. S., & Whelan, D. C. (2012). *Would introverts be better off if they acted more like extraverts? Exploring emotional and cognitive consequences of counterdispositional behavior.* *Emotion (Washington, D.C.),* 12(2), 290–303. [https://doi.org/10.1037/a0025169](https://doi.org/10.1037/a0025169)

Zettler, I., Hilbig, B. E., & Heydasch, T. (2013). Two sides of one coin: Honesty–humility and situational factors mutually shape social dilemma decision making. *Journal of Research in Personality*, 47(4), 286–295. [https://doi.org/10.1016/j.jrp.2013.01.012](https://doi.org/10.1016/j.jrp.2013.01.012)

Ziegler, M., Kemper, C. J., & Kruyen, P. (2014). Short scales – five misunderstandings and ways to overcome them. *Journal of Individual Differences*, 35(4), 185–189. [https://doi.org/10.1027/1614-0001/a00148](https://doi.org/10.1027/1614-0001/a00148)
SUPPLEMENTARY MATERIALS

Peer Review History
Download: https://collabra.scholasticahq.com/article/27386-be-yourself-and-behave-appropriately-exploring-associations-between-incongruent-personality-states-and-positive-affect-tiredness-and-cognitive-perf/attachment/68275.docx?auth_token=z2_RKdw2kssarfl-r-oN

Deviations from the Preregistration
Download: https://collabra.scholasticahq.com/article/27386-be-yourself-and-behave-appropriately-exploring-associations-between-incongruent-personality-states-and-positive-affect-tiredness-and-cognitive-perf/attachment/68276.pdf?auth_token=z2_RKdw2kssarfl-r-oN

Supplemental Materials
Download: https://collabra.scholasticahq.com/article/27386-be-yourself-and-behave-appropriately-exploring-associations-between-incongruent-personality-states-and-positive-affect-tiredness-and-cognitive-perf/attachment/68277.html?auth_token=z2_RKdw2kssarfl-r-oN