INTERNATIONAL REVIEW OF SOCIAL PSYCHOLOGY

Adelina, N., & Feldman, G. (2021). Are Past and Future Selves Perceived Differently from Present Self? Replication and Extension of Pronin and Ross (2006) Temporal Differences in Trait Self-Ascription. *International Review of Social Psychology*, 34(1): 29, 1–16. DOI: https://doi.org/10.5334/irsp.571

RESEARCH ARTICLE

Are Past and Future Selves Perceived Differently from Present Self? Replication and Extension of Pronin and Ross (2006) Temporal Differences in Trait Self-Ascription

Nadia Adelina and Gilad Feldman

We attempted a pre-registered replication and extension of Studies 1, 2, and 3 from Pronin and Ross (2006) regarding the effects of social and temporal distance on trait attributions with an online American Amazon MTurk sample ($N = 911$). We concluded mixed findings. We found support for the original findings: participants attributed more dispositional traits to others compared to themselves, although with weaker effects (original: $f = 0.35$, 95% CI [0.09, 0.61]; replication: $f = 0.10$, 95% CI [0.03, 0.16]). Also, similar to the original, we found that participants tended to attribute a favorable ratio of positive traits when making self-assessments (original: $f = 0.77$, 95% CI [0.29, 1.25]; replication: $f = 0.88$, 95% CI [0.50, 1.26]). However, unlike the original, we failed to find support for the core hypothesis that participants would ascribe more dispositional traits to their temporally distant self compared to their present self (original: $f = 0.54$, 95% CI [0.27, 0.77]; replication: $f = 0.02$, 95% CI [0.00, 0.06]). Furthermore, in contrast to the original, we found that the positive traits ratio increases with temporal distance (original: $f = 0.16$, 95% CI [0.00, 0.36]; replication: $f = 0.33$, 95% CI [0.22, 0.42] in the opposite direction). Contrary to our hypothesis, in an extension, we found that people were more likely to ascribe a greater ratio of positive traits to their friends than to themselves ($\xi = 0.3$, 95% CI [0.21, 0.38]). All materials, data, and code are provided here: https://osf.io/gs2rx/.

Keywords: temporal asymmetry; temporal distance; bias; judgment and decision making; psychological distance; actor-observer asymmetry; trait ascriptions

Introduction

We see ourselves differently than how we see others (Pronin, 2008). This is partly related to our limited ability to read others' minds. We have access to our internal thoughts and feelings, yet we can only make assessments about other people based on their observable behaviors. An example of this process is a phenomenon psychologists refer to as the *actor-observer asymmetry in attribution* (Jones & Nisbett, 1971), wherein people are more likely to make dispositional attributions when assessing others but make situational inferences when assessing themselves.

Expanding on this, Pronin and Ross (2006) posited that a similar asymmetry applies to the judgments we make about our temporally distant self. Unlike the assessments we make about ourselves in the present, how we judge our past and future self is more akin to the way we judge others. By asking participants to make trait assessments about their friends versus themselves in either the past or present (Study 1), their present or future selves (Study 2), and their past, present, or future selves (Study 3), Pronin and Ross (2006) were able to demonstrate that the trait assessments we make about our temporally distant selves more closely resemble those we make about other people's traits than those we typically make about ourselves. Importantly, these findings were later linked with reasons for why people often make poor long-term choices (e.g., Ersner-Hershfield, Wimmer, & Knutson, 2009; Jones, Hine, & Marks, 2017). As Pronin and Ross (2006) would argue, it may be because we see our temporally distant self 'as an other'.

In light of the academic impact and practical implications of Pronin and Ross (2006), we sought to conduct a very close replication of their studies with two clear goals. The first goal was to conduct an independent close replication of the temporal asymmetry phenomenon. The second goal was to examine an extension regarding the effect of self-serving bias on trait self-ascriptions.

We begin by introducing the literature on psychological distance as it relates to the phenomena of actor-observer

Department of Psychology, University of Hong Kong, Hong Kong SAR, HK

Corresponding author: Gilad Feldman (gfeldman@hku.hk)

asymmetry and temporal asymmetry in trait ascriptions. Following that, we introduce the target article, Pronin and Ross (2006), highlighting motivation for the current replication study and providing an overview of the scope of the replication. We then summarize the original hypotheses and the findings by Pronin and Ross (2006) and suggest an extension examining self-serving bias.

Psychological Distance

Psychological distance refers to the subjective perceived distance of how far away or, conversely, how close something is in regard to its social, temporal, spatial, or hypothetical proximity from the self in the present moment (Liberman et al., 2007; Trope & Liberman, 2010). Although we are only able to directly experience the present moment, construal level theory (CLT; Trope & Liberman, 2010) postulates that we are able to transcend the present moment and traverse psychological distances by creating abstract mental construals of psychologically distant objects. It is through this process of creating mental construals that people are able to do things like reminisce about the past (temporal distance), plan for the future (temporal distance), and empathize with points of view other than one's own (social distance) despite those being outside the bounds of their direct experience. Since these different kinds of psychological distances all operate with the self as the reference point, Trope and Liberman (2010) suggested that they operate using similar cognitive mechanisms and thus have similar effects on perception and action. Since Pronin and Ross (2006) investigated the comparable effect of temporal distance and social distance on trait ascriptions, we will be focusing on these facets of psychological distance for the present discussion.

According to CLT, the more psychologically distant an object is from the self, the more abstract the construal becomes. Meaning, higher-level construals of more distant objects tend to be more abstract and general (e.g., excel in academics), whereas lower-level construals of less distant objects are more concrete and detailed (e.g., read a textbook). These construals go on to affect one's judgment or perception of said object. For example, studies have found that people make increasingly dispositional attributions, as opposed to situational attributions, about other actors' behaviors as temporal distance increases (e.g., Funder & Van Ness, 1983). This is because higher-level construals promote the tendency to interpret distant behaviors in terms of decontextualized and abstract dispositions rather than concrete, situational factors (Trope & Liberman, 2010). As will be elaborated in the following paragraphs, this can affect practical processes such as moral judgment (e.g., Agerström & Björklund, 2009; Mårtensson, 2017) and risk-taking behaviors (e.g., Hershfield & Kramer, 2017; Raue et al., 2015).

Social Distance: Actor-Observer Asymmetry in Trait Assessments

The actor-observer asymmetry refers to the observed difference in the type of assessments we make regarding other people compared to those we make of ourselves. Whereas people often make dispositional attributions for the actions that other people make (e.g., 'She did not give up her seat for the elderly lady on the train because she is rude.'), they often resist doing so for their own actions, instead attributing them to situational factors (e.g., 'I did not give up my seat for the elderly lady because of my sprained ankle.'). In their seminal paper regarding actor-observer asymmetry, Jones and Nisbett (1971) argued that this phenomenon results from a difference in the information available to the actors versus those available to the observers. Although we are aware of the situational constraints influencing our own behavior (e.g., a sprained ankle), this information is not as salient to observers. Thus, in the absence of available information, observers attribute the behaviors of other actors to their disposition. Similarly, according to CLT, this is because other people are more psychologically distant than the self and therefore construed on a higher-level. Hence, when making judgments about a temporally distant self, we tend to focus on superordinate, dispositional traits rather than subordinate, situational explanations.

Temporal Distance: Temporal Asymmetry in Trait Assessments About the Self

A similar asymmetry in trait ascriptions can also be observed when comparing assessments made of the present self versus those made of temporally distant selves. As Pronin and Ross (2006) found, participants were more inclined to attribute dispositional traits, and less inclined to make ascriptions to situational variability, to their past or future self compared to their present self. This may be because temporal distance changes people's mental representations of themself (Trope & Liberman, 2003). More specifically, the greater the temporal distance, the more abstract the mental construal becomes, focusing on the perceived essence of their identity (e.g., personality) rather than specific details (e.g., situational constraints). As a result, people ascribe a greater number of dispositional traits to their temporally distant selves compared to their present self.

The theories and studies discussed thus far relate closely to Pronin and Ross's (2006) study, in which they extended upon the existing literature by positing that the temporal asymmetry in trait self-ascriptions closely mirrors the classic actor-observer asymmetry in trait ascriptions. Participants are more likely to ascribe dispositional traits, instead of situational ascriptions, to others compared to themselves. Similarly, participants are also more inclined to ascribe dispositional traits to their past and future selves compared to their present self. As previously discussed, this is because increases in both temporal and social distance cause higher-level construals. Thus, they may have comparable effects, possibly with similar underlying mechanisms. In this way, they argued that the temporally distant self may be perceived 'as an other'.

Choice of Study for Replication

We chose to replicate Pronin and Ross (2006) based on two factors: absence of direct replications and impact. Although the article has been highly influential, to our knowledge, there are no published direct replications of this work. At the time of writing, there were a total of 298 Google Scholar citations of the article and many impor-

tant follow-up theoretical and empirical articles. Although Pronin and Ross (2006) did not investigate the effects of temporal and social distance on measures beyond trait attributions, high-impact follow-up research has shown the implications of temporal distance on behaviors relating to moral actions (e.g., Agerström & Björklund, 2009; Hershfield, Cohen & Thompson, 2012; Van Gelder, Hershfield & Nordgren, 2013) and long-term saving (Bryan & Hershfield, 2013; Ersner-Hershfield et al., 2009).

These studies extended Pronin and Ross's (2006) findings and demonstrated practical implications. For example, when making assessments about the ethicality of morally questionable behavior (e.g., choosing to not participate in blood donation during a health crisis), people are more likely to give harsher moral judgments when the behavior is described to occur in the distant future as opposed to the near future (Agerström & Björklund, 2009). Relating back to CLT, because distant future events are construed on a higher level, people are more likely to attribute morally questionable behaviors to disposition (e.g., 'She's selfish.') as opposed to situational factors (e.g., 'She just recovered from a flu and thus cannot donate blood right now.'), leading to harsher judgments.

Furthermore, although prior researchers have suggested parallels in temporal asymmetry for both past and future selves (e.g., Trope & Liberman, 2003), Pronin and Ross (2006) were the first to directly compare assessments made across multiple temporal distances (past vs. present vs. future) and social distances (self vs. friend). By doing so, they were able to expand upon the literature by demonstrating that manipulations in temporal distance have a similar effect to manipulations in social distance, as predicted by CLT.

Pronin and Ross (2006) has been a highly influential article. A meta-analysis by Malle (2006) of 173 published studies regarding the actor-observer asymmetry surprisingly found only very weak effects ($d = 0.016$ to $d = 0.095$), the significance of which depended on the specifics of the study's design (e.g., intimacy of the relationship between actor and observer, how variables were coded, how hypothetical events were described, and valence of hypothetical events). These findings called into question the robustness of the classic actor-observer bias, especially in regard to its applicability to real-world situations which are highly idiosyncratic. Considering the similarity between the effects of temporal distance and social distance on trait attributions, this may indicate similar concerns regarding the replicability and robustness of temporal distance effects. As discussed by Coles et al. (2018), in situations wherein the reliability of findings are uncertain, direct replications may be of greater utility than conceptual replications. Furthermore, the small sample size of the original study may indicate it was underpowered and with possibly overestimated effect sizes.

Given these reasons, we aimed to revisit the classic phenomenon to examine the reproducibility and replicability of the findings by conducting an independent replication of the study. Following the recent growing recognition of the importance of reproducibility and replicability in psychological science (e.g., Open Science Collaboration, 2015; Zwaan et al., 2018), we embarked on a well-powered, pre-registered very close replication of Pronin and Ross (2006). Also, by doing so, we aim to contribute to obtaining a more precise estimate of the observed effects for this phenomenon.

Summary of Original Studies and Current Replication

Overall, Pronin and Ross (2006) consisted of seven studies examining the effect of temporal distance on trait ascriptions. We focused the current replication on Studies 1, 2, and 3 due to their similar hypotheses and experimental design. In this section, we outline the hypotheses (refer to Table S5 in supplementary), experimental design (refer to Table S1 in supplementary), and findings (refer to Table S6 in supplementary) of each study before providing an overview of the current replication based on these studies. We report a more comprehensive analysis of each study in the section 'Analysis of Studies 1–3 from Pronin and Ross (2006)' in the supplementary.

In Study 1, Pronin and Ross (2006) tested the hypothesis that participants would ascribe a greater number of traits, and a lesser number of ascriptions to situational variability when assessing their friends (Hypothesis 1) or their past selves (Hypothesis 2) compared to their present selves. Confirming their hypothesis, they found that compared to an assessment of their present selves, participants indeed ascribed a greater number of traits to their friends ($f = 0.35$ [0.09, 0.61]) and their past selves ($f = 0.43$ [0.17, 0.69]).

Expanding on Study 1, in Study 2, the authors demonstrated that this temporal asymmetry in trait ascriptions not only applies to those made of the past self but also for the future self. Participants did ascribe more traits to their future selves compared to their present selves ($f = 0.51$ [0.17, 0.84]).

Lastly, in Study 3, the authors investigated an alternate hypothesis for the observed temporal asymmetry in self trait ascriptions detected in Studies 1 and 2—the self-enhancement hypothesis. According to this alternate hypothesis, participants ascribe a greater number of traits to their past (Study 1) and future self (Study 2) due to a self-enhancing motive to see themselves as improving over time, not due to the manipulation of temporal distance. If this were true, participants would ascribe more negative traits to their past self, more ascriptions to situational variability for their present self, and more positive traits to their future self.

As such, they first sought to replicate the findings of Studies 1 and 2 that participants would ascribe a greater number of traits to their past and future selves compared to their present selves. Further, they hypothesized that if self-enhancement was the true motive behind temporal asymmetry, then participants would ascribe more negative traits to their past selves, more ascriptions to situational variability for their present selves, and more positive traits for their future selves. They replicated the findings from Study 1 and 2 ($f = 0.54$ [0.27, 0.77]), yet failed to find support for the self-enhancement hypothesis. Although participants attributed more positive than negative traits to their present selves ($f = 0.77$ [0.29, 1.25]), the effect was weaker for temporal distance ($f = 0.16$ [0.00, 0.36]).

Overview of Current Replication

Given the similarities between Studies 1, 2, and 3, we combined the experimental design of the three studies into a single 3 (temporal distance: past vs. present vs. future) × 2 (social distance: self vs. friend) between-subjects experimental design for the current replication. We summarized details regarding the adjustments to the current replication in comparison to the original in **Table 1**.

Extension: Self-Serving Bias in Trait Ascriptions About the Self

As an extension to the replication study, we wanted to examine whether, as a result of self-serving bias (SSB), participants would ascribe a greater ratio of positive traits compared to negative traits when they were making judgments about themselves in contrast to when they were making judgments about others.

In Study 3 of Pronin and Ross (2006), the authors found that regardless of temporal distance, participants ascribed a greater number of positive traits compared to negative traits when making judgments about themselves, taken to demonstrate SSB. However, they did not investigate whether the same applies when participants make judgments about others (e.g., their friends). By combining the design of Study 1, which examined differences in situational and trait attributions for self vs. friend judgments, and Study 3, which measured differences in negative and positive trait attributions for self-judgments across temporal distance, we were able to compare the differences in the ratio of positive and negative trait attributions across both temporal and social distance simultaneously.

According to the SSB, people are more likely to attribute personal failures to situational factors (e.g., 'I failed the test *because I was sick* the night before and was not able

Table 1: Comparison of original versus the replication: Adjustments and reasons for change.

Item	Original	Replication	Reason for change
Study design	3 separate designs: Study 1: 2 (past vs. present) × 2 (self vs. friend) between-subjects design Study 2: present self vs. future self, between-subjects design Study 3: mixed design. Past vs. present vs. future self between subjects design, and negative vs. positive trait attributions as within subjects	3 (temporal distance: past vs. present vs. future) × 2 (social distance: self vs. friend) between-subjects design	Combined Studies 1–3 into one study due to similar experimental designs that do not conflict with each other
Conditions	Participants randomly assigned to one of ... · Four experimental conditions (past friend vs. present friend vs. past self vs. future self) in Study 1	Participants randomly assigned to one of six experimental conditions: · 6 between-subject experimental conditions in total: past-self vs. present self vs. future self vs. past friend vs. present friend vs. future friend · (Added future friend condition as an extension)	Same as above
Procedure	Participants were told to answer the scales according to the condition they were randomly assigned to. (RA asked to verify manipulation for IV2)	Same instructions as the original with a few additions: · Added comprehension check throughout the survey · Added manipulation check for IV2 (positive vs. negative traits) at the end of the survey	To account for the MTurk sample, we had to add comprehension checks to make sure participants were paying attention. A manipulation check was also added, since in the original study they had an RA verify the manipulation, but it may not be applicable for MTurk sample.
Measures	Online web survey: · Study 1 and Study 2 used the same scales. Study 3 used a modified scale with opposite valenced traits. · 1 measure has 11 scales. · Items are in fixed order.	Online Qualtrics survey: · Combined scales from Studies 1 and 2 and Study 3 into one set of scales with the order of the items randomized. · 1 measure has 22 scales.	Since we combined the experimental design of Studies 1–3, we combined the two measures in one, in randomized order, as separating them would have statistical implications.
Participant population	Undergraduate university students and university staff (only Study 1) recruited through a university database email	Recruited through the online platform Amazon MTurk	To make the results more generalizable beyond university populations. Since MTurk samples are more diverse agewise, we did not have to separate them into two sample groups (university staff and university students)

to study.'), yet attribute others' failures to dispositional factors (e.g., 'She failed the test *because she's lazy.*'). This is differentiated from the actor-observer bias because the SSB specially applies to valenced situations, wherein people are likely to attribute negative incidents to situational explanations and positive incidents to dispositional factors when assessing themselves but not others. The reasons for the SSB are twofold: motivational factors (i.e., to self-enhance and present the self in a favorable light) and differences in the availability of information regarding one's own actions versus other's actions (Shepperd et al., 2008). The meta-analysis by Malle (2006) supports this idea, finding that the actor-observer asymmetry is stronger for negative events, whereas the opposite occurs for positive events.

Drawing from this theory of SSB, we therefore hypothesized that participants would attribute fewer negative traits when making assessments about themselves but not others. We predicted that, in general, participants would ascribe fewer negative traits to themselves, as they would explain away negative experiences as resulting from situational variables instead of negative dispositional factors. In contrast, positive experiences are more likely to be attributed to positive dispositional factors instead of situational variables. Furthermore, since SSB results from motivational factors and the disparity in self vs. other information, we predicted that this effect would be greater when making self-judgments compared to judgments of others.

Practically, these findings may have important implications. Although SSB can be adaptive when applied in certain circumstances, it can also be maladaptive in others. For example, in situations which cannot be changed, attributing these negative outcomes to internal causes can have negative consequences on one's self-esteem, sense of agency, and affect regulation. In these cases, attributing to external factors is more optimal. However, in situations where change is possible, attributing negative outcomes to external causes can prevent people from improving their outcomes. By repeatedly attributing negative events (e.g., financial, academic, or relationship failures) to situational factors (e.g., bad luck, fault of others), people may fail to learn from bad experiences. Take the example of a bankrupt investor who is repeatedly investing in the same type of bad businesses and losing increasingly more money or of a person who has been through multiple divorces and is repeatedly entering romantic relationships with incompatible partners. By evading personal responsibility, they fail to learn and make amendments to their behaviors to prevent similar events from happening in the future, as they never see it as being 'their fault'. Evidently, a better understanding of SSB processes is crucial in regard to enabling us to maximize its adaptive advantages while minimizing its negative consequences.

Pre-Registration and Open Science

We first pre-registered the experiment on the Open Science Framework (OSF), and data collection was launched later that week (https://osf.io/yrvuq). Datasets and R/RMarkdown code were made available on the OSF (https://osf.io/gs2rx/). Open science details and disclosures, power analyses, and all materials used in these experiments are available in the supplementary materials. All measures, manipulations, and exclusions conducted for this investigation are reported; all studies were pre-registered with power analyses reported in the supplementary; and data collection was completed before analyses.

Method

Power Analysis

To ensure that the current replication sample has sufficient power, we calculated effect sizes (ES) and power based on the statistics reported in the target article. To ensure that the sample can detect even the smallest effect, we chose the smallest yet still statistically significant ($p < 0.05$) observed effect size (calculated in Table S13 in the supplementary) for each study to base our power analyses on.

Furthermore, since publication bias often inflates ESs detected in published studies (Ioannidis, 2008), and considering the heterogeneity of ES across studies (McShane & Bockenholt, 2014), replications based on the reported ES of the target study may be underpowered (Shrout & Rodgers, 2018). To combat this, we implemented the *safeguard power analysis method* (Perugini, Gallucci, & Costantini, 2014), which bases power analysis on the lower 60% confidence interval of the target effect size. This method involved calculating the 60% confidence interval (e.g., 0.19, 0.45) of the target effect size and then basing the power analysis on the lower bound effect size (e.g., 0.19). Afterwards, we then chose the maximum required sample size out of the three to become the required sample for the current replication. We concluded that the minimum required sample size for a power of 0.95 and alpha of 0.05 is 362 participants. We provided more information regarding these calculations in Tables S13 and S14 in the supplementary.

Participants and Sensitivity Analysis

Based on these aforementioned calculations, we recruited a total of 911 American Amazon Mechanical Turk (MTurk) using CloudResearch/TurkPrime (Litman et al., 2016) in return for USD 0.63 based on calculations of USD 7.5/hour to meet minimum federal wage. After excluding the participants who fulfilled the pre-registered exclusion criteria (refer to the 'Pre-Exclusions Versus Post-Exclusions' section in the supplementary), this sample was reduced to 878. However, it should be noted that the replication sample ($n = 291$) is smaller than the estimated sample for Study 1 ($n = 362$). We conducted a post hoc sensitivity analysis using the pwr.t.test() function from the *pwr* (Champely et al., 2018) package in R and found that the replication was sensitive to detect an effect of $d = 0.39$ with 95% power and effect of 0.29 with 80% power (one-sided; refer to 'VI. Sensitivity Analysis' in the supplementary for further details). A comparison of the target article sample and the replication samples is provided in **Table 2**.

Design and Procedure

We summarized the experimental design in Table S15 in the supplementary. Based on our analysis of the original

Table 2: Difference and similarities between original study and replication (after exclusions).

	Pronin and Ross (2006)				American MTurk workers		
	Study 1		Study 2	Study 3	Self conditions	Other conditions	Overall
	Student sample	Staff sample					
Sample size	123	44	40	75	438	440	878
Geographic origin	–	–	–	–	US American		
Gender	42 males, 81 females	65 females, 3 males	26 females, 14 males	39 females, 35 males, 1 not reported	232 females, 203 males, 3 not disclosed	224 females, 210 males, 6 not disclosed	413 males, 456 females, 9 not disclosed
Median age (years)	19	47	–	–	37	36	37
Average age (years)	–	–	–	–	40.09	39.4	39.75
Age range (years)	–	–	–	–	18–73	19–76	18–76
Medium (location)	Completed via a web-based questionnaire						
Compensation	Course requirement	None	Course credit	Candy bars	Monetary		
Year	2006				2020	2020	2020

article (see 'Analysis of Studies 1–3 from Pronin & Ross (2006)' in supplementary), we decided to combine the experimental design of Studies 1, 2, and 3 from Pronin and Ross (2006) into a single 3 (temporal distance: past and present and future) by 2 (social distance: self and friend) between-subject design due to their methodological similarity and non-conflicting design (refer to the corresponding section in the summary for full justification on the combined design). The display of conditions was counterbalanced. We provided more details and all measures in the 'Materials and Scales Used in the Replication + Extension Experiment' section in the supplementary.

Participants were randomly assigned to one of six experimental conditions, and they responded to a series of scales. At the end of the experiment, they answered a number of funneling questions and provided their demographic information. We provided a more comprehensive overview of the survey procedure in Table S17 in the supplementary.

Manipulations

Each participant was randomly assigned to provide assessments about one of the following targets: (1) self in the past, (2) self in the present, (3) self in the future, (4) a friend in the past, (5) a friend in the present, or (6) a friend in the future. We provided additional details of the differences in manipulation between the six conditions, the experimental design, and the complete scales used in the current replication in Table S15 in the supplementary.

IV1: Temporal Distance (Between)

We asked participants to make judgments about either themselves or their friend (depending on IV2) 'five years ago' (past), 'right now' (present), or 'five years from now' (future).

IV2: Social Distance (Between)

We asked participants to assess either themselves or a friend that they have known for at least five years.

Measures

Total Trait Ascriptions

Differences in trait ascriptions were measured using a set of 22 items. This set was created by combining the two scales used in Pronin and Ross (2006) Studies 1–3 into a single set, which were then presented in randomized order to the participants to address any potential order effects. Each scale presented three possible options: two opposing traits (e.g., serious-carefree) and a third option (variable/depends on the situation). For each scale, participants had to choose one of the three options which best describe the person designated by the experimental condition they were assigned to. For example, if they were assigned to the 'past self' condition, they had to choose which of the three options best described themselves five years ago. Similarly, if they were assigned to the 'future friend' condition, they had to choose which would best describe their friend five years from now. Afterwards, each participant's answers were converted into a single score representing the total amount of trait ascriptions (e.g., choosing 'serious' or 'carefree') they made as opposed to ascriptions to situational variability (e.g., choosing 'variable/depends on the situation' instead of 'serious' or 'carefree'). In the supplementary, we detailed the 22 items presented to the participants in Table S15 and the original scales used in Studies 1, 2, and 3 in Pronin and Ross (2006) in Table S4.

Negative and Positive Trait Ascriptions

To test the self-enhancement hypothesis, Pronin and Ross (2006) modified the 11 scales in Study 3 so that the trait pairs in each scale were oppositely valenced, meaning that

one was obviously positive (P) while the other was obviously negative (N). In the present study, we combined the 11 scales used in Study 3 with the 11 scales used in Studies 1 and 2, resulting in a total of 22 scales. As detailed in Table S15, scales numbered 12 to 22 were oppositely valenced to the first 11 scales. For example, the pair 'serious-carefree' was modified to become 'uptight (P)–easygoing (N)'. Each participant's answers were then summarized into two scores, one representing the total number of positive trait ascriptions they made (as opposed to negative trait ascriptions or ascriptions to situational variability) and the other representing the total number of negative trait ascriptions they made.

Valence of Traits

To ensure that the manipulation of valenced trait pairs were effective, participants were asked to rate 22 traits from the 11 scales in regard to how desirable or desirable each one is on a Likert scale of –3 (*Very negative*) to 3 (*Very positive*).

Evaluation Criteria for Replication Findings and Classification

To evaluate the results of the current replication, we followed the recommendations set by LeBel, Vanpaemel, Cheung, and Campbell (2019) and did so based on the following criteria: signal presence, consistency between the effect size of the original study and the present replication, and the precision of its estimate between the replication and target study. Results of the present replication will be interpreted following the aforementioned criteria, with consistency referring to the effect size instead of the direction of the effect. We provided more details in Figure S2 under the section 'Evaluation Criteria for Replication Findings' in the supplementary.

To classify the replication, we followed the criteria by LeBel, McCarthy, Earp, Elson, and Vanpaemel (2018) which evaluates the methodological similarity between the original study and the replication study depending on various design facets (refer to Figure S2 in the supplementary for specific details). Based on this, we classify the current replication as a 'very close replication' (refer to **Table 3** for specific details).

Pre-Registered Data Analysis Strategy

To reduce 'researcher degrees of freedom' (Simmons et al., 2011), we carried out data analysis following our pre-registered plan (https://osf.io/yrvuq). Any changes that we made from this pre-registered plan were noted and justified in the 'Comparisons and Deviations' section in the supplementary. We used R/RStudio (R Core Team, 2013) to complete data cleaning, manipulation, and analyses. For a detailed discussion of how we detected and managed outliers and assumptions of the statistical tests, refer to the corresponding sections in the supplementary.

Table 3: Classification of the replication, based on LeBel et al. (2018).

Design facet	Replication	Details of deviation
IV operationalization	Same	–
DV operationalization	Same	–
IV stimuli	Same	–
DV stimuli	Similar	The two different sets of scales (each containing 11 scales) from Studies 1–2 and 3, respectively, were combined into one set of scales (containing 22 scales) presented in randomized order.
Procedural details	Similar	1. For the friend conditions, we asked participants to note the initials of the friend they were thinking of so it can be piped into the following questions. 2. We added comprehension checks to ensure that participants were paying attention. 3. We added a manipulation check for IV2 at the end of the survey which asked individuals to rate 22 traits in terms of their valence (how desirable/undesirable).
Physical settings	Similar	Both studies were conducted via a web-based questionnaire. The present study uses Qualtrics and the Amazon MTurk platform.
Contextual variables	Different	The original sample was predominantly undergraduate university students and female. Meanwhile, the present sample is recruited from the Amazon MTurk platform.
Replication classification	Very close replication	

Note: See supplementary Figure S3 and Lebel et al. (2018) for details about categorization and the added 'Similar' rational, referring to 'minor deviations or extensions aimed to adjust the study to the target sample that are not expected to have major implications on replication success'.

Effect of Temporal and Social Distance on Trait Ascriptions

In each of the three studies reported in Pronin and Ross (2006), they conducted a between-subjects analysis of variance (ANOVA) with temporal distance (past vs. present in Study 1; present vs. future in Study 2; past vs. present vs. future in Study 3) as the IV and total trait ascriptions as the DV to test the effect of temporal distance on trait ascriptions. Additionally, in Study 1, they also conducted a between-subjects ANOVA with social distance (self vs. friend) as the IV and trait ascriptions as the DV. In Study 3, they also added a second DV: the total number of positive trait ascriptions.

Given that we combined the designs of Studies 1, 2, and 3 in the current replication, we began our analysis by conducting a 3 (past vs. present vs. future) × 2 (self vs. friend) between-subjects multivariate analysis of variance (MANOVA), with total number of trait ascriptions (DV1) and total number of positive trait ascriptions (DV2) as the dependent variables, to examine whether temporal distance and social distance have an impact on trait ascriptions. This choice was based on the recommendations by Huberty and Morris (1992). A MANOVA was used to examine how both the total number of positive trait ascriptions (DV2) and the number of trait ascriptions (DV1) accounted for the group differences between the experimental conditions and also to investigate how variations in temporal distance and social distance relatively contributed to the group differences in outcome variables within a multivariate context. This would be beyond the scope of multiple ANOVAs.

Afterwards, we conducted follow-up ANOVAs for each DV and IV, respectively, to further examine the underlying factors contributing to the MANOVA results we had obtained. Given that these variables were initially studied in univariate contexts in the studies by Pronin and Ross (2006) (refer to the original article analysis in supplementary for more details), we conducted multiple ANOVAs to enable better comparison with the original studies and to examine the individual effects.

As such, to first investigate whether the total number of trait ascriptions (DV1) varied based on temporal distance (IV1) and social distance (IV2), we conducted a 2 (self vs. other) × 3 (past vs. present vs. future) between-subjects ANOVA with total traits as the DV. We also did the same for the total number of positive traits (DV2). Then, to examine whether social distance (IV2) had an impact on trait ascriptions (DV1), we conducted a between-subjects ANOVA with social distance (self vs. other) as the independent variable and the total number of trait ascriptions as the DV. Then, to examine whether temporal distance (IV1) influenced trait ascriptions (DV1), we conducted another between-subjects ANOVA with the total number of trait ascriptions as the DV but with temporal distance (past vs. present vs. future) as the independent variable.

Finally, to replicate the findings of Studies 1 and 2 of Pronin and Ross (2006), we conducted two independent t-tests, one comparing past vs. present self trait ascriptions (Study 1), and the other comparing present vs. future self trait ascriptions (Study 2).

Self-Enhancement Hypothesis

To test the alternative self-enhancement hypothesis in Study 3, Pronin and Ross (2006) first conducted a dependent t-test comparing negative and positive trait ascriptions in the present-self group followed by a one-way ANOVA comparing the ratio of positive-to-total trait ascriptions in the past-, present-, and future-self conditions. Following the original study, we first conducted a paired samples t-test to see whether there are differences in the ratio of positive trait ascriptions in the present-self group, taken as demonstrating a motive to self-enhance. Then, to examine whether this ratio varied across different temporal differences in the self conditions, we conducted a one-way ANOVA (past vs. present vs. future) with the ratio of positive-to-total trait ascriptions as the DV. If we found support, we followed with post hoc tests in the form of an independent samples t-test to further explore the differences.

Extension: Positivity Effect in Trait Attributions About the Self

As an extension, we conducted a one-way MANOVA with social distance (self vs. friend) as a predictor and total negative traits and total positive traits as the dependent variables. We did so to investigate whether people would ascribe a greater number of positive traits and a lesser number of negative traits when making assessments about themselves compared to when they were making assessments about others. If we found support for the model, we would follow with two one-way ANOVA for negative traits and positive traits separately.

Results

Replication

We summarized all descriptive statistics of all measures in **Table 4**. Statistical tests of the hypotheses are plotted in **Figure 1** for total trait ascriptions (DV1) and **Figure 2** for total number of positive trait ascriptions (DV2). We summarized a comparison of the findings of the current replication and the original findings of Studies 1, 2, and 3 from Pronin and Ross (2006) in **Table 5**.

Outlier Detection and Assumptions Checks

Following the pre-registered plan, we first screened the data for both multivariate and univariate outliers. We included more information regarding the results of these assumption checks in Table S20 in the supplementary. Since the tests revealed that the data did not meet the assumptions of normality and skewness, we conducted the planned tests using robust statistical methods as a supplemental analysis in addition to the parametric versions to mirror what was originally done by Pronin and Ross (2006). We found no indication for differences between the results of the parametric and the robust tests, and we therefore focused our reporting of the results on the parametric tests. The results of the robust tests are presented in Table S21 in the supplementary.

Table 4: Descriptive statistics.

	IV1: Past *M(SD)*			IV1: Present *M(SD)*			IV1: Future *M(SD)*			Overall *M(SD)*		
	Total traits	**Positive traits**	**Negative traits**	**Total traits**	**Positive traits**	**Negative traits**	**Total traits**	**Positive traits**	**Negative traits**	**Total traits**	**Positive traits**	**Negative traits**
IV2: Self	16.17 (4.95)	5.03 (2.62)	3.18 (2.43)	16.28 (4.52)	5.64 (2.70)	2.58 (2.20)	16.03 (4.17)	6.86 (2.43)	1.49 (1.76)	16.16 (4.55)	5.84 (2.69)	2.42 (2.25)
		N = 147			*N* = 144			*N* = 147			*N* = 438	
IV2: Friend	17.23 (4.21)	6.68 (2.54)	2.17 (2.17)	16.63 (3.92)	6.74 (2.51)	1.98 (2.07)	17.21 (3.61)	7.20 (2.39)	1.70 (1.83)	17.02 (3.92)	6.87 (2.48)	1.95 (2.03)
		N = 146			*N* = 147			*N* = 147			*N* = 440	
Overall	16.70 (4.62)	5.85 (2.71)	2.68 (2.35)	16.46 (4.23)	6.20 (2.65)	2.28 (2.15)	16.62 (3.94)	7.03 (2.41)	1.60 (1.79)	16.59 (4.27)	6.36 (2.64)	2.18 (2.16)
		N = 293			*N* = 291			*N* = 294			*N* = 878	

Note: M indicates mean. *SD* indicates standard deviation. *N* indicates the sample size for that box. Total traits were calculated by averaging the total number of trait ascriptions (choosing one of the two trait descriptors and not choosing the third situational descriptor) each participant made in each condition. As discussed in the manipulations section, positive/negative were only in 11 out of the 22 items. As such, the range of total traits is 0–22, while the range for positive and negative is 0–11 (referring only to items 11–22 that had valence).

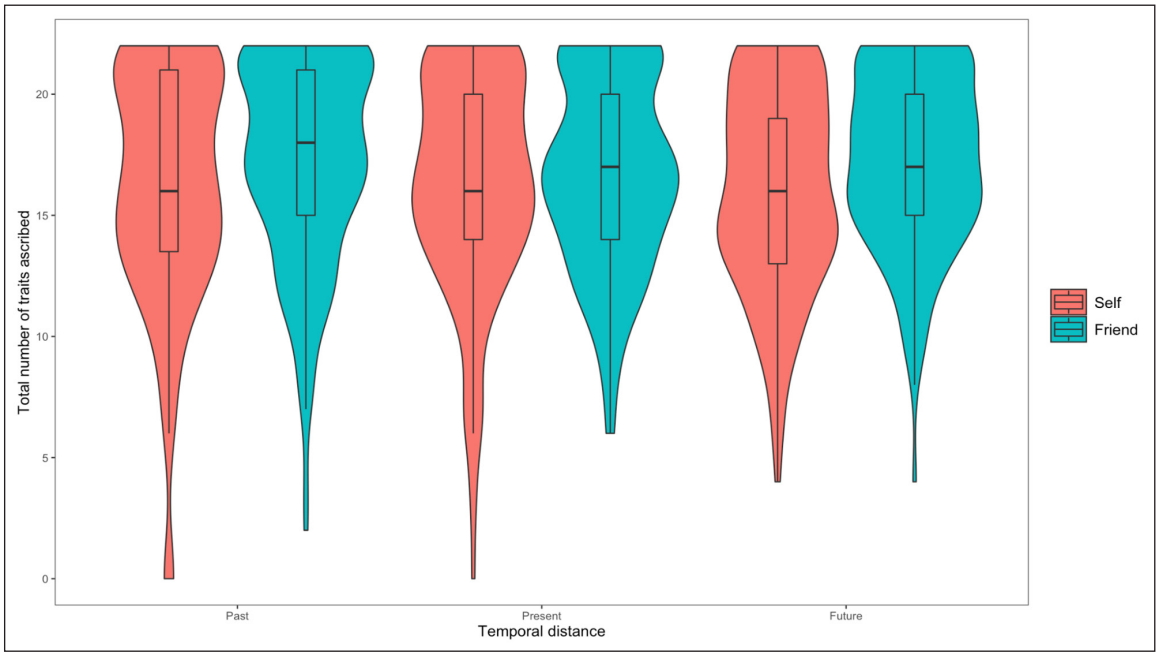


Figure 1: Plot for the mean number of overall trait ascriptions participants made regarding themselves versus others across temporal distance with a possible range of 0–22.

Effects of Temporal and Social Distance on Total Trait Ascriptions and Positive Trait Ascriptions

To investigate the effects of social and temporal distance on trait ascriptions, we first conducted a 3 (temporal: past vs. present vs. future) × 2 (social: self vs. other) between-subjects MANOVA. Similar to what Pronin and Ross (2006) found in their Study 1, we found support for multivariate main effects of temporal distance ($F(4, 1742) = 10.65$, $p < 0.001$, $f = 0.16$, 95% CI [0.10, 0.20]). Additionally, we found support for multivariate main effects of social distance (F (2, 871) = 18.17, $p < 0.001$, $f = 0.20$, 95% CI [0.13, 0.27]), and the interaction between temporal distance and social distance (F (4, 1742) = 3.89, $p = 0.004$, $f = 0.09$, 95% CI [0.03, 0.13]) on the number of participants' trait ascriptions. This suggests that participants varied in the type of trait ascriptions they made (situational vs. dispositional), depending on social and temporal distance.

We conducted two additional follow-up 2 × 3 ANOVAs on each DV to further tease apart the effect of both temporal and social distance on the number of trait ascriptions (DV1) and the total number of positive trait ascriptions (DV2), and we found support for the effect of social distance on both DV1 and DV2 yet failed to find support for temporal distance.[1] In the follow-up analyses presented below, we further explored these differences.

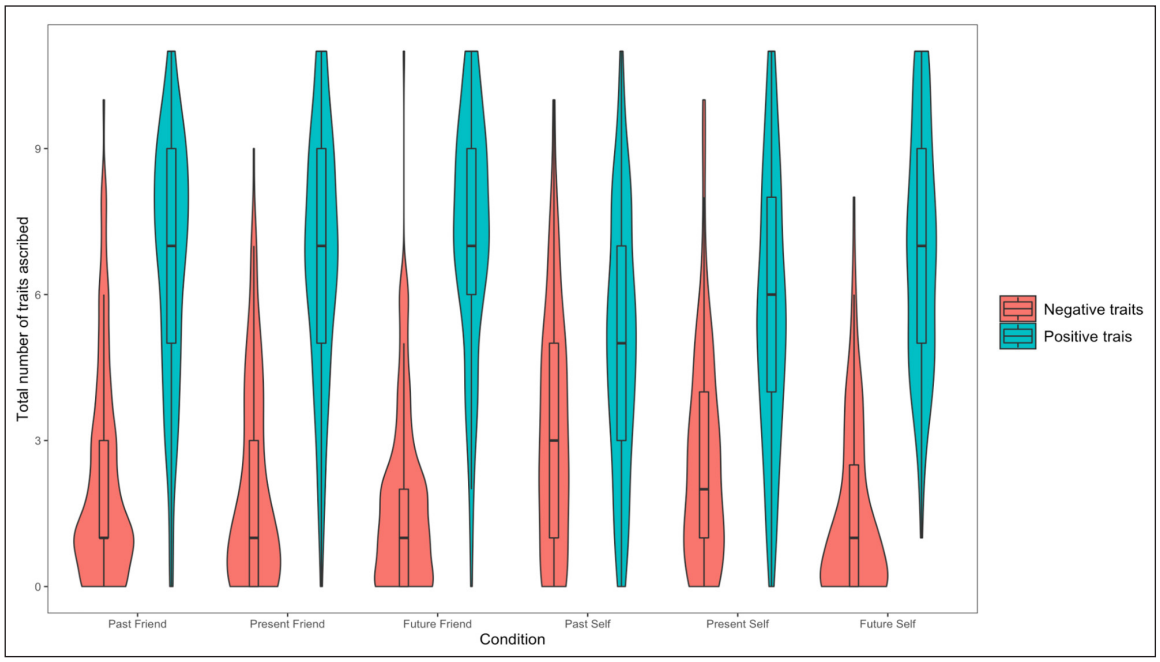


Figure 2: Plot for the mean number of overall negative and positive trait ascriptions in each experimental condition, with a possible range of 0–11 (referring only to items 11–22 that had valence).

Table 5: Summary and comparison of findings of the current replication study and those of Pronin and Ross (2006) based on the criteria by LeBel et al. (2019).

Study	Hypothesis		Target effect	Original effect size (Cohen's *f*)	Replication effect size (Cohen's *f*)	Interpretation
	No.	Description				
1	1a	Actor-observer asymmetry in self vs. other (friend) trait ascriptions	Social distance (self vs. friend)	0.35, 95% CI [0.09, 0.61]	0.10, 95% CI [0.04, 0.17]	Signal Inconsistent Smaller
	1b	Temporal asymmetry resembles actor-observer asymmetry	Temporal distance (past vs. present)	0.43, 95% CI [0.17, 0.69]	–0.01, 95% C.I. [–0.21, 0.19]	No signal Inconsistent
2	2	Temporal asymmetry in trait self-ascriptions	Temporal distance (present vs. future)	0.51, 95% CI [0.17, 0.84]	0.03, 95% C.I. [–0.20, 0.25]	No signal Inconsistent
3	3a	Self-enhancement hypothesis of temporal asymmetry in trait self-ascriptions	Ratio of positive-to-total trait ascriptions (present self)	0.77, 95% CI [0.29, 1.25]	0.88, 95% CI [0.50, 1.26]	Signal Consistent
	3b		Ratio of positive-to-total trait ascriptions (past vs. present vs. future self)	0.16, 95% CI [0.00, 0.36]	0.33, 95% CI [0.22, 0.42]	Signal Inconsistent Larger
	3c	Temporal asymmetry	Temporal distance (past vs. present vs. future)	0.54, 95% CI [0.27, 0.77]	0.02, 95% CI [0.00, 0.06]	No signal Inconsistent

Social Distance: Self-Other Asymmetry in Total Trait Ascriptions

Mirroring Pronin and Ross's (2006) Study 1, we conducted a follow-up univariate ANOVA to further investigate the effect of social distance on trait ascriptions. We found support for the hypothesis that trait ascriptions would differ based on social distance, $F(1, 876) = 9.02$, $p = 0.003$, $f = 0.10$, 95% CI [0.04, 0.17]. That is, participants ascribed a greater number of traits, as opposed to ascriptions to situational variability, when making judgments about their friend ($M = 17.02$, $SD = 3.92$, $N = 440$) compared to when they were making judgments about themselves ($M = 16.16$, $SD = 4.55$, $N = 438$). Since Cohen's f CI excluded zero but did not include the original ES point estimate, we concluded that a signal was detected, yet the replication ES for hypothesis 1a was inconsist-

ent with the original, as the magnitude of the effect was smaller.

Temporal Distance: Temporal Asymmetry in Total Trait Ascriptions About the Self

Then, to further investigate the effect of temporal distance on participants' trait ascriptions, we conducted another univariate ANOVA aiming to replicate the findings in Pronin and Ross's (2006) Study 3, followed by independent t-tests aiming to replicate the findings of their Studies 1 and 2. Unlike the findings of Study 3, we failed to find support for the hypothesis that trait ascriptions would differ based on temporal distance, $F(2, 872) = 0.24$, $p = 0.791$, $f = 0.02$, 95% CI [0.00, 0.06]. Since Cohen's f CI included zero and excluded the original ES point estimate, we concluded that no signal was detected and that the replication ES for hypothesis 3c is inconsistent with the original.

Inconsistent with the findings of Study 1 from Pronin and Ross (2006), we failed to find support for differences between the trait ascriptions that participants made of their past ($M = 16.17$, $SD = 4.95$, $N = 147$) or present self ($M = 16.28$, $SD = 4.52$, $N = 144$), $t(287.58) = -0.21$, $p = 0.837$, $f = -0.01$, 95% CI [–0.21, 0.19]. Since Cohen's f CI included zero and excludes the original ES point estimate, we conclude that no signal was detected and that the replication ES for hypothesis 1b is inconsistent with the original.

Furthermore, inconsistent with the findings in Pronin and Ross's (2006) Study 2, we failed to find support for differences in the trait ascriptions people made between their present self ($M = 16.63$, $SD = 3.92$, $N = 144$) and future self ($M = 17.21$, $SD = 3.61$, $N = 147$), $t(286.01) = 0.49$, $p = 0.623$, $f = 0.03$, 95% CI [–0.20, 0.25]. Since Cohen's f CI included zero and excluded the original ES point estimate, we concluded that no signal was detected and that the replication ES for hypothesis 2 is inconsistent with the original ES.

Self-Enhancement Hypothesis: Temporal Asymmetry in Ratio of Positive-to-Total Trait Ascriptions

Present Self: Ratio of Positive-to-Total Trait Ascriptions

To test the self-enhancement hypothesis, we first conducted a paired-samples t-test. In accordance with the findings from Study 3 in Pronin and Ross (2006), participants were more likely to ascribe positive traits ($M = 5.64$, $SD = 2.70$, $N = 144$) to their present self compared to negative traits ($M = 2.58$, $SD = 2.20$, $N = 144$), $t(274.72) = 10.55$, $p < 0.001$, $f = 0.88$, 95% CI [0.50, 1.26]. Since Cohen's f CI excluded zero and included the original ES point estimate, we concluded that signal was detected and that the replication ES for hypothesis 3a was consistent with the original.

Temporal Distance and Self: Ratio of Positive-to-Total Trait Ascriptions

We then proceeded to investigate whether the ratio of positive-to-total trait ascriptions would vary over different temporal distances using a one-way ANOVA. In contrast to the findings of Study 3 in Pronin and Ross (2006), we found that participants attributed a different ratio of positive-to-total traits depending on temporal distance, $F(2, 430) = 22.83$, $p < 0.001$, $f = 0.33$, 95% CI [0.22, 0.42]. More specifically, participants ascribed a greater ratio of positive traits in the future-self condition ($M = 0.44$, $SD = 0.15$, $N = 147$) compared to the present-self condition ($M = 0.35$, $SD = 0.15$, $N = 144$) and past-self condition ($M = 0.32$, $SD = 0.17$, $N = 147$). Since Cohen's f CI excluded zero and excluded the original ES point estimate, we concluded that the replication ES for hypothesis 3b was inconsistent with the original, as the magnitude of the effect was larger.

Extension: Self-Other Asymmetry in Positive Trait Ascriptions

We ran an extension to examine whether people would ascribe more positive traits and fewer negative traits to themselves compared to their friends. Since data screening revealed that the data did not meet the assumptions required for parametric tests, we conducted a robust one-way MANOVA to test this. We found that the number of positive and negative trait ascriptions varied depending on social distance, $F(2, 875) = 34.76$, $p < 0.001$.

Following this, we conducted robust one-way ANOVAs to investigate the contribution of each dependent variable to the main effects. For the model with positive traits as the dependent variable, we found that participants were more inclined to ascribe a greater number of positive traits to their friends ($M = 6.87$, $SD = 2.48$, $N = 440$) compared to themselves ($M = 5.84$, $SD = 2.69$, $N = 438$), $F(1, 525.1) = 48.86$, $p < 0.001$, $\xi = 0.3$, 95% CI [0.21, 0.39]. For the model with negative traits as the dependent variable, participants were more likely to ascribe a greater number of negative traits when they were making judgments about themselves ($M = 2.42$, $SD = 2.25$, $N = 438$) compared to those they made of their friends' ($M = 1.95$, $SD = 2.03$, $N = 440$) conditions, $F(1, 524.19) = 11.00$, $p < 0.001$, $\xi = 0.17$, 95% CI [0.07, 0.26].

In summary, participants surprisingly ascribed a more favorable ratio of positive traits compared to negative traits to their friends than to themselves.

Comparing Replication to Original Findings

We compared the target article to our replication using the LeBel et al. (2019) replication evaluation criteria and summarized our findings in **Table 5**. Whenever we detected a signal in the replication, it was always in the same direction as in the original. Interestingly, although hypothesis 3b was not supported in the original study, the present replication was able to find support for the hypothesis, consistent with the original authors' reasoning. Detection of the effect was possibly due to our replication being better powered than the original.

Discussion

We conducted a pre-registered replication of temporal asymmetry (Pronin & Ross, 2006) with the twofold aim of assessing the replicability of its findings and extending the study by investigating the impact of social distance on the self-enhancement hypothesis. The evaluation of the current replication study was done according to the

criteria set by LeBel et al. (2019), and the corresponding comparison of the results of the target article and the current study is presented in **Table 5**. We summarized information regarding the similarities and differences between the original article and the present replication in the 'Original Versus Replication' section in the supplementary.

Overall, the findings of the current replication were inconclusive, as it was mixed in terms of consistency with the original findings of Pronin and Ross (2006) regarding the effect of social distance, temporal distance, and the motive to self-enhance. As will be further discussed, we successfully replicated some effects, although we were unable to find support for the core hypotheses regarding temporal asymmetry. For the effect of social distance, we found that, although smaller in the magnitude of the effect, participants were more likely to ascribe dispositional traits, as opposed to situational ascriptions, when making assessments about their friends compared to themselves. We were unable to replicate the effect of temporal distance, as no signal was detected, and the ES is inconsistent with the original. Whereas Studies 1, 2, and 3 of Pronin and Ross (2006) found that participants attributed a greater number of dispositional traits to their past or future self compared to their present self, the current study was not able to detect such a difference. However, we were able to find support for the alternate self-enhancement hypothesis, as the ES detected signal and is consistent with the original. Meaning, participants tended to attribute a favorable ratio of positive-to-total traits when making assessments about themselves in the present. Furthermore, whereas the original did not, we found that this ratio increased with temporal distance. Participants ascribed the greatest number of positive traits when making assessments about their future self compared to their present and past self, in respective order, possibly reflecting a motive to self-enhance. Although Pronin and Ross (2006) hypothesized this in Study 3, they failed to find support for it. Meanwhile, the present study did, with an ES that detected a signal and was inconsistent and larger than the original. These mixed findings will be further elaborated in the following paragraphs.

Replication

Social Distance

Although smaller than the magnitude of the original effect, we were able to detect a signal for the effect of social distance on trait ascriptions (refer to Table 10 for comparison). One reason for the smaller effect may be due to the small sample size of the original study (refer to the original article analysis in the supplementary). This may have resulted in an inflated effect that may not be representative of the true effect size (Ioannidis, 2008). This is in line with the results of a recent meta-analysis regarding actor-observer asymmetry by Malle (2006), who analyzed 173 published studies on the actor-observer asymmetry in attributions and found much smaller effects ($d = 0.016$ to $d = 0.095$) than anticipated. As such, this replication contributes to the emerging evidence, which may suggest a need for a re-examination of the strength of the effect of social distance on attribution. More specifically, despite the ubiquity of this effect in various textbooks and research, the strength of this effect may not be as strong as would be suggested by the published literature.

Temporal Distance

We were unable to replicate the effect of temporal distance on trait ascriptions. We were unable to detect a signal, and the effect was considerably weaker than in the original (see **Table 5**). We believe that it is unlikely that the failure to replicate this effect resulted from a methodological differences and that the findings may suggest that the effect of temporal distance may be weaker in reality and possibly applies mainly to valenced traits, suggesting a need to reframe the hypotheses for temporal distance.

Although we combined the designs of Studies 1, 2, and 3 into a single experimental design for the current replication, we argue that this does not interfere with the closeness or quality of the replication. Based on our analysis and comparison of the three studies (refer to the corresponding section in the supplementary under 'Comparisons and Deviations'), the designs did not conflict with one another when combined, mostly due to the methodological similarity and between-subjects design. Otherwise, the method of the present replication is very close to the original, as we used the same list of traits as the experimental stimuli and conducted the experiment using a web-based questionnaire. Furthermore, although combining the studies into one may cause concerns in regard to effect size and power, we addressed this through using the safeguard power analysis method (Perugini, Gallucci, & Costantini, 2014), ensuring that the samples for the tests for each of the hypotheses was sufficiently powered.

Hence, given that the effect of social distance demonstrated in a meta-analysis (Malle, 2006) was smaller than initially thought and the similarity between the processes by which social distance and temporal distance are theorized to affect trait judgements via construals, it is possible that the effect of temporal distance is also not as strong as would be suggested by the published literature. Additionally, the findings of the present study suggest that temporal distance does affect trait ascriptions, but only for valenced traits. Nonetheless, future studies or replications will be required to confirm our findings and obtain a more precise estimate of the observed effects for this phenomenon.

Self-Enhancement Hypothesis

Similar to the original, we found that participants ascribed a greater ratio of positive-to-total traits to themselves. In contrast to the original, we found that this ratio increased with temporal distance (refer to Table 10 for comparison). That is, participants seemed to demonstrate a self-serving bias to see themselves as improving over time and developing in increasingly positive ways. Malle (2006) found that the actor-observer asymmetry only emerged for negative events, whereas the opposite occurred for posi-

tive events. However, controlling for valence, this difference disappeared. Thus, this may be taken to indicate that the support for an actor-observer asymmetry in attributions may not be as robust as initially thought and that the documented actor-observer effects in the existing literature may result not from a person-situation distinction but rather from a self-serving bias, or perhaps a different explanation altogether. Our findings must be further tested and confirmed by future studies which can investigate how variations in valence, scenarios, or trait descriptions can influence the actor-observer effect.

Extension

We ran an extension examining the difference in positive versus negative trait attributions between self and other attributions, hypothesizing that participants would ascribe a greater ratio of positive traits when making assessments about themselves compared to their friends out of a motive to self-enhance and see themselves in a positive light. This was based on prior studies which found that people tend to attribute more positive traits during self-appraisal as opposed to other-appraisal out of a motivation to enhance self-worth (e.g., Brown, 1986; Steele, Spencer, & Lynch, 1993). Surprisingly, the results of the current investigation suggest that the opposite may be true. That is, people were more likely to ascribe positive traits when making judgments about their friends compared to when they were assessing themselves.

One possible reason for this unexpected finding is that we had participants make judgments of a close friend who they have known for five years, rather than an acquaintance or stranger. Many of the studies regarding the actor-observer asymmetry have been done in the context of non-intimate relationships (e.g., making judgments about new acquaintances or strangers) rather than on intimate relationships (making judgments of parents or child, close friends, or romantic partners) (Malle, 2006). In this regard, Jones and Nisbett (1971) postulated that relational intimacy functioned as a moderator of actor-observer differences, wherein people making judgments about others with whom they are in an intimate relationship with would demonstrate a less salient difference compared to strangers or acquaintances who they do not know well. As such, future studies can investigate how this effect may vary as a function of relational intimacy: for example, whether participants are asked to make judgments about new acquaintances or strangers instead of close friends.

Implications, Limitations, and Directions for Future Research

Constraints on Generality (COG)

Participants

Since exclusively American participants were recruited using MTurk, there may be concerns regarding the demographic representativeness of the current sample. This has been raised in previous investigations using such samples (e.g., Walters, Christakis, & Wright, 2018, Huff & Tingley, 2015). As such, this may limit the generalizability of the present findings to other populations without these characteristics, and it is uncertain how these findings would apply to non-WEIRD samples (Western, educated, industrialized, rich, democratic; Henrich, Heine, & Norenzayan, 2010) which were not accounted for in the present sample. This is an important limitation, as prior studies (e.g., Choi, Nisbett, & Norenzayan, 1999) have found that the tendency towards dispositional and situational attributions of self versus others vary cross-culturally. More specifically, they found that East Asians tend to demonstrate less correspondence bias than Western samples due to the importance placed on situational context on behaviors in Asian thinking. As such, it is noted that our ability to generalize is limited, and it is recommended that future research explore these potential cross-cultural differences with greater specificity.

Materials/Stimuli

Following Pronin and Ross (2006), we used a list of descriptive trait pairs as the experimental stimuli. Since this was a limited number of traits, it is possible that using a list with a different variety or category of traits may lead to different results, especially because Malle (2006) found that the type of traits (e.g., external, internal, positive, negative) influenced the magnitude of the actor-observer asymmetry in trait judgments. For example, Malle (2006) found that participants showed stronger actor-observer asymmetry in judgment of intimates for external attributions but not internal attributions. For this reason, studies must be careful to include manipulation checks which capture the type of traits used as experimental material.

Procedures

Malle (2006) has shown that research design (between-subject vs. within-subject) had an effect on the difference in trait attributions for self versus others. Differences in judgments of self compared to others only emerged in between-subject designs but not within-subject designs. Since the present study used a between-subject design, it is possible that the observed effects may not occur in a within-subjects design.

Constraints on Theory Generalizability

We observed an effect for social distance and self-enhancement but not for temporal distance, and we concluded mixed support for the findings in the original article. Given the link suggested between the phenomena and these experimental designs and findings, we interpreted this to be in support of the phenomena in this specific context and methodology. We note that our ability to generalize from these findings to other contexts and methodology is limited and implications for theory need to be further elaborated and tested.

Conclusion

Our findings were inconclusive, with mixed results regarding the replicability of Pronin and Ross (2006). We were able to successfully replicate the effects of social distance and self-enhancement on trait attributions. However, we were unable to replicate the central

hypothesis regarding the effect of temporal distance on trait ascriptions in the present replication, suggesting the need for more replications to test our findings and obtain a more precise estimate regarding the effect of this phenomenon.

Note

[1] For the model with total trait ascriptions as the DV, we found support for social distance predicting differences in trait ascriptions, $F(1, 872) = 9.02$, $p = 0.003$, $f = 0.10$, 95% CI [0.05, 0.16], but not temporal distance, $F(2, 872) = 0.24$, $p = 0.791$, $f = 0.02$, 95% CI [0.00, 0.06]. Participants ascribed more traits to their friends than themselves, $p = 0.003$. As for the model with the total number of positive trait ascriptions as the DV, we found that participants' likelihood of ascribing positive traits varied based on both temporal distance, $F(2, 872) = 16.88$, $p < 0.001$, $f = 0.20$, 95% CI [0.14, 0.25], and social distance, $F(1, 872) = 36.20$, $p = 0.003$, $f = 0.20$, 95% CI [0.15, 0.26]. Post hoc comparisons showed differences in the number of positive traits in the future versus past conditions, $p < 0.001$ as well as in the future versus present conditions, $p < 0.001$, but with no support for the present versus past conditions, $p = 0.225$.

Additional File

The additional file for this article can be found as follows:

- **Supplementary materials.** Pronin and Ross (2006) Replication and Extension: Supplementary. DOI: https://doi.org/10.5334/irsp.571.s1

Competing Interests

The authors have no competing interests to declare.

Author Contributions

Nadia conducted the replication as part of her undergraduate thesis in psychology. Gilad was the thesis advisor. He supervised each step in the project, conducted the pre-registrations, and ran data collection. Nadia and Gilad jointly finalized editing the thesis for journal submission.

Author Information

Nadia Adelina is a graduate student with the University of Hong Kong psychology department.

Gilad Feldman is an assistant professor with the University of Hong Kong psychology department. His research focuses on judgment and decision-making.

Additional Information

The current replication is part of the larger 'mass pre-registered replications in judgment and decision-making' project. The project aims to revisit well-known research findings in the area of judgment and decision making (JDM) and to investigate the replicability of these findings.

Contributor Roles Taxonomy

The table below employs CRediT (Contributor Roles Taxonomy) to identify the contributions and roles of the contributors in the current replication effort. Refer to https://www.casrai.org/credit.html for details and definitions of each of the roles listed below.

Role	Nadia Adelina	Gilad Feldman
Conceptualization		X
Pre-registration	X	X
Data curation	X	
Formal analysis	X	
Funding acquisition		X
Investigation	X	
Pre-registration peer review/verification		X
Data analysis peer review/verification		
Methodology	X	
Project administration		X
Resources		X
Software	X	
Supervision		X
Validation	X	
Visualization	X	
Writing, original draft	X	
Writing, review and editing	X	X

References

Agerström, J., & **Björklund, F.** (2009). Temporal distance and moral concerns: Future morally questionable behavior is perceived as more wrong and evokes stronger prosocial intentions. *Basic and Applied Social Psychology*, *31*(1), 49–59. DOI: https://doi.org/10.1080/01973530802659885

Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition*, *4*(4), 353–376. DOI: https://doi.org/10.1037/2325-9965.1.S.2

Bryan, C. J., & **Hershfield, H. E.** (2013). You owe it to yourself: Boosting retirement saving with a responsibility-based appeal. *Decision*, *1*(S), 2–7. DOI: https://doi.org/10.1037/2325-9965.1.S.2

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., ... & **De Rosario, M. H.** (2018). Package 'pwr'. *R package version*, *1*(2).

Choi, I., Nisbett, R. E., & **Norenzayan, A.** (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*(1), 47. DOI: https://doi.org/10.1037/0033-2909.125.1.47

Coles, N., Tiokhin, L., Scheel, A. M., Isager, P. M., & **Lakens, D.** (2018). The costs and benefits of replication studies. *Behavioral and Brain Sciences*, *41*. DOI: https://doi.org/10.1017/s0140525x18000596

Ersner-Hershfield, H., Wimmer, G. E., & **Knutson, B.** (2009). Saving for the future self: Neural measures of future self-continuity predict temporal discounting. *Social Cognitive and Affective Neuroscience*, *4*(1), 85–92. DOI: https://doi.org/10.1093/scan/nsn042

Funder, D. C., & **Van Ness, M. J.** (1983). On the nature and accuracy of attributions that change over time. *Journal of Personality*, *51*(1), 17–33. DOI: https://doi.org/10.1111/j.1467-6494.1983.tb00851.x

Henrich, J., Heine, S. J., & **Norenzayan, A.** (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29. DOI: https://doi.org/10.1038/466029a

Hershfield, H. E., Cohen, T. R., & **Thompson, L.** (2012). Short horizons and tempting situations: Lack of continuity to our future selves leads to unethical decision making and behavior. *Organizational Behavior and Human Decision Processes*, *117*(2), 298–310. DOI: https://doi.org/10.1016/j.obhdp.2011.11.002

Hershfield, H. E., & **Kramer, L. A.** (2017). Examining the effect of social distance on financial decision-making. In *Work Pap*. Available from https://rady.ucsd.edu/docs/seminars/kramer-paper.pdf [Google Scholar].

Huberty, C. J., & **Morris, J. D.** (1992). Multivariate analysis versus multiple univariate analyses. DOI: https://doi.org/10.1037/10109-030

Huff, C., & **Tingley, D.** (2015). "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics*, *2*(3). DOI: https://doi.org/10.1177/2053168015604648

Ioannidis, J. P. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19*(5), 640–648. DOI: https://doi.org/10.1097/EDE.0b013e31818131e7

Jones, C., Hine, D. W., & **Marks, A. D.** (2017). The future is now: Reducing psychological distance to increase public engagement with climate change. *Risk Analysis*, *37*(2), 331–341. DOI: https://doi.org/10.1111/risa.12601

Jones, E. E., & **Nisbett, R. E.** (1971). The actor and the observer: Divergent perceptions of the causes of behavior. General Learning Press.

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & **Vanpaemel, W.** (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, *1*, 389–402. DOI: https://doi.org/10.1177/2515245918787489

LeBel, E. P., Vanpaemel, W., Cheung, I., & **Campbell, L.** (2019). A brief guide to evaluate replications. *Meta-Psychology*, *3*, 1–9. DOI: https://doi.org/10.15626/MP.2018.843

Liberman, N., Trope, Y., & **Stephan, E.** (2007). Psychological distance. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (pp. 353–381). Guilford Press.

Malle, B. F. (2006). How the mind explains behavior: Folk explanations, meaning, and social interaction. MIT Press.

Mårtensson, E. (2017). Construal level theory and moral judgments: How thinking abstractly modifies morality. *Journal of European Psychology Students*, *8*(1), 30–40. DOI: https://doi.org/10.5334/jeps.413

McShane, B. B., & **Böckenholt, U.** (2014). You cannot step into the same river twice: When power analyses are optimistic. *Perspectives on Psychological Science*, *9*(6), 612–625. DOI: https://doi.org/10.1177/1745691614548513

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). DOI: https://doi.org/10.1126/science.aac4716

Perugini, M., Gallucci, M., & **Costantini, G.** (2014). Safeguard power as a protection against imprecise power estimates. *Perspectives on Psychological Science*, *9*(3), 319–332. DOI: https://doi.org/10.1177/1745691614528519

Pronin, E. (2008). How we see ourselves and how we see others. *Science*, *320*(5880), 1177–1180. DOI: https://doi.org/10.1126/science.1154199

Pronin, E., & **Ross, L.** (2006). Temporal differences in trait self-ascription: When the self is seen as an other. *Journal of Personality and Social Psychology*, *90*(2), 197. DOI: https://doi.org/10.1037/0022-3514.90.2.197

Raue, M., Streicher, B., Lermer, E., & **Frey, D.** (2015). How far does it feel? Construal level and decisions under risk. *Journal of Applied Research in Memory and Cognition*, *4*(3), 256–264. DOI: https://doi.org/10.1016/j.jarmac.2014.09.005

Shepperd, J., Malone, W., & **Sweeny, K.** (2008). Exploring causes of the self-serving bias. *Social and Personality Psychology Compass*, *2*(2), 895–908. DOI: https://doi.org/10.1111/j.1751-9004.2008.00078.x

Shrout, P. E., & **Rodgers, J. L.** (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*, 487–510. DOI: https://doi.org/10.1146/annurev-psych-122216-011845

Simmons, J. P., Nelson, L. D., & **Simonsohn, U.** (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. DOI: https://doi.org/10.1177/0956797611417632

Steele, C. M., Spencer, S. J., & **Lynch, M.** (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology*, *64*(6), 885. DOI: https://doi.org/10.1037/0022-3514.64.6.885

Trope, Y., & **Liberman, N.** (2003). Temporal construal. *Psychological Review*, *110*(3), 403. DOI: https://doi.org/10.1037/0033-295X.110.3.403

Trope, Y., & **Liberman, N.** (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440. DOI: https://doi.org/10.1037/a0018963

Van Gelder, J. L., Hershfield, H. E., & **Nordgren, L. F.** (2013). Vividness of the future self predicts

delinquency. *Psychological Science*, *24*(6), 974–980. DOI: https://doi.org/10.1177/0956797612465197

Walters, K., Christakis, D. A., & **Wright, D. R.** (2018). Are Mechanical Turk worker samples representative of health status and health behaviors in the US? *PloS One*, *13*(6), e0198835. DOI: https://doi.org/10.1371/journal.pone.0198835

Zwaan, R. A., Etz, A., Lucas, R. E., & **Donnellan, M. B.** (2018). Making replication mainstream. *Behavioral and Brain Sciences*, *41*. DOI: https://doi.org/10.1017/S0140525X17001972

How to cite this article: Adelina, N., & Feldman, G. (2021). Are Past and Future Selves Perceived Differently from Present Self? Replication and Extension of Pronin and Ross (2006) Temporal Differences in Trait Self-Ascription. *International Review of Social Psychology*, 34(1): 29, 1–16. DOI: https://doi.org/10.5334/irsp.571

Submitted: 27 January 2021 **Accepted:** 27 September 2021 **Published:** 21 December 2021

]u[*International Review of Social Psychology* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS

Pronin and Ross (2006) replication and extension: Supplementary

Contents

Open Science disclosures

Data and code

Data and code are shared using the Open Science Framework. Review link for data and code of the study: https://osf.io/gs2rx/

Pre-registrations and Qualtrics study designs

Link: https://osf.io/yrvuq

Procedure and data disclosures

Data collection

Data collection was completed before analyzing the data.

Conditions reporting

All collected conditions are reported.

Data exclusions

Details are reported in the materials section of this document

Variables reporting

All variables collected for this study are reported and included in the provided data.

Analysis of studies 1-3 from Pronin & Ross (2006)

I.) Original Article Methods

A.) Experimental Design

Please see Table S1 for an overview of the experimental design of Studies 1-3, and Table S4 to view screenshots of the design as described by the authors in the original article.

Table S1

Overview of studies 1-3 from Pronin & Ross (2006)

Study	Independent variable (IV)	Dependent variable (DV)	Experimental Design
1	IV1: Temporal distance (past vs present) IV2: Social distance (self vs friend)	DV1: Attributional style (actor-like vs observer-like)	2 (social distance) x 2 (temporal distance) between-subject design
2	IV1: Temporal distance (present vs future)	DV1: Attributional style (actor-like vs observer-like)	between-subject design
3	IV1: Temporal distance (past vs present vs future)	DV1: Attributional style (actor-like vs observer-like) DV2: Attribution of positive and negative traits	between-subject design

B.) Independent variables

In study 1, each participant was randomly assigned to one of four conditions, wherein they were asked to provide assessments about either : 1.) themselves in the past, 2.) themselves in the present, 3.) their friend in the past, **or** 4.) their friend in the present on a set of 11 scales measuring actor-observer differences in trait attributions. Please see table S2 for details of each condition in study 1, and table S3 to view the scales.

It should be noted that two samples were involved in study 1. First, a larger sample of undergraduate students and second, a smaller sample of university staff. The university staff responded to the same scales (DV), but were only asked to provide assessments of themselves either in the present or 10 years ago. This was done to generalize the results beyond college-aged individuals (see table S10, description of results from university staff sample).

In study 2, participants followed the same procedure and responded using the same measure as study 1, but were randomly assigned to one of two conditions (instead of four in study 1) and asked to provide assessments of either: 1.) themselves in the present ("5 years ago") or 2.) themselves in the future ("5 years from now").

In study 3, participants were randomly assigned to one of three conditions and asked to provide assessments of themselves in either : 1.) the past ("5 years ago"), 2.) the present, or 3.) the future ("5 years from now") on a modified version of the 11 scales used in studies 1 and 2.

Table S2

Experimental manipulation of independent variables in study 1 of Pronin & Ross (2006)

IV2: Social distance (self vs friend)	**IV1: Temporal distance (past vs present)**	
	IV1: Past (between)	**IV1: Present (between)**
IV2: Self (between)	1.) Past-Self Condition Manipulation : participants asked to provide assessments about themselves "5 years ago"	2.) Present-Self Condition Manipulation : participants asked to provide assessments about themselves in the present
IV2 : Friend (between)	3.) Past-Friend Condition Manipulation : participants asked to provide assessments about a friend "5 years ago" who	4.) Present-Friend Condition Manipulation : participants asked to provide assessments about a friend in

they had been friends with for five years, and who they were still friends with.

the present who they have been friends with for five years.

C.) Dependent variables

Please see table S4 for an overview of the dependent variables across the three studies, including how they were measured.

DV1: Attributional style (AS). For studies 1 and 2, AS was assessed on 11 scales (Nisbett et al., 1973) measuring actor-observer differences in trait ascriptions (see table S3). Each scale contained three possible options: two opposing traits (e.g. serious-carefree) and a third option (variable/ depends on the situation). For each scale, participants had to choose one of the three options which best describes, depending on the condition they were assigned to, themselves/their friend in the past/present/future. Each participant's responses were then converted into a single score representing the total number of trait ascriptions they made (number of times participant chooses either trait options), as opposed to attribution to situational variability (number of times participant chooses variable/depends on the situation). Figure S1 displays screenshots from Nisbett et al. (1973) describing the original scales, which were later adapted by Pronin & Ross (2006) for their study.

DV2: Attribution of positive and negative traits (APN). In study 3, both AS and APN was assessed on a modified version of the 11 scales used in studies 1 and 2, wherein the pairs of traits for each scale was changed slightly so that one was obviously positive (P) while the other was obviously negative (N) (e.g. serious-carefree became uptight (N)–easygoing (P)). Participants' scores were summarized into two scores, one measuring the total number of trait attributions they made (measures AS), and another measuring the total number of positive trait attributions they made (measures APN).

TABLE 4

POLAR TRAIT TERMS USED IN STUDY III

Polar trait term	
Serious–*gay*	Skeptical–trusting
Subjective–*analytic*	Quiet–talkative
Future oriented–present oriented	Cultivated–*natural*
Energetic–relaxed	*Sensitive*–tough-minded
Unassuming–self-asserting	Self-sufficient–sociable
Lenient–firm	Steady–*flexible*
Reserved–*emotionally expressive*	Dominant–deferential
Dignified–*casual*	Cautious–*bold*
Realistic–idealistic	Uninhibited–self-controlled
Intense–*calm*	*Conscientious*–happy-go-lucky

Note. Where trait terms differ in social desirability at the .05 level or more, the trait higher in social desirability is italicized.

Procedure

Twenty-four subjects filled out questionnaires indicating, for themselves and four other stimulus persons, which of three descriptions best fit the stimulus person: a trait term, its polar opposite, or the phrase "depends on the situation." Subjects responded to a total of 20 such three-choice items for each of the stimulus persons. Questionnaire booklets were arranged so that for the group of subjects as a whole, the questionnaire for each stimulus person preceded the questionnaire for every other stimulus person equally often. On a final questionnaire, subjects rated the desirability of each of the 40 polar traits on a 7-point scale (−3 to +3).

Figure S1. The original scales from Nisbett et al. (1973, p.161) along with its instructions.

Table S3

Overview of dependent measures used in studies 1-3 of Pronin & Ross (2006)

Study	Dependent variable (DV)	Specific Measure
1 2	**DV1: Attributional style (observer-like vs actor-like)** Specific DV item: number of trait ascriptions made (compared to attributions to situational variability). ● **Actor-like** : less trait ascriptions ● **Observer-like** : more trait ascriptions	**Scale:** adapted from 11 scales measuring observer-actor difference in trait ascriptions from Nisbett et al. (1973) ● serious– carefree—variable/depends on the situation ● subjective–analytic—variable/depends on the situation ● energetic–relaxed—variable/depends on the situation ● unassuming–self-asserting—variable/depends on the situation ● lenient–firm—variable/depends on the situation ● intense– calm—variable/depends on the situation ● quiet–talkative—variable/depends on the situation ● introverted– extroverted—variable/depends on the situation ● steady–flexible—variable/depends on the situation ● cautious– bold—variable/depends on the situation ● cooperative– competitive —variable/depends on the situation
3	**DV1: Attributional style (observer-like vs actor-like)** Specific DV item: same as with studies 1 and 2 **DV2: Attribution of positive vs negative traits (self enhancement hypothesis)** Specific DV item: ratio of the number of positive traits (denoted by "(P)", as opposed to "(N)" for negative traits) to total traits selected.	**Scale:** modified from 11 scales measuring observer-actor difference in trait ascriptions from Nisbett et al. (1973) ● uptight (N)–easygoing (P)–variable/depends on the situation ● fickle (N)– reasonable (P)–variable/depends on the situation ● energetic (P)–lazy (N)–variable/depends on the situation ● shy (N)–self-assured (P)–variable/depends on the situation ● passive (N)–decisive (P)–variable/depends on the situation ● frenzied (N)–cool-headed (P)–variable/depends on the situation ● concise (P)–wordy (N)–variable/depends on the situation ● reclusive (N)–sociable (P)–variable/depends on the situation ● stubborn (N)–adaptable (P)–variable/depends on the situation ● timid (N)–brave (P)–variable/depends on the situation

- helpful (P)–selfish (N)–variable/depends on the situation

Table S4

Screenshots of the method and procedure sections of each of the three studies.

Study	Method and procedure as described in Pronin & Ross (2006)	Page no.
1	*Procedure and dependent measure.* The undergraduate participants were randomly assigned to provide assessments regarding themselves "in the present," themselves "5 years ago," a friend in the present with whom they had been friends for at least 5 years, or a friend "5 years ago" with whom they had been friends for at least 5 years and with whom they were still friends. The assessments were made on a series of 11 scales developed by Nisbett et al. (1973) in an investigation of actor–observer differences in the attribution process.[2] Each scale featured two opposing character traits (i.e., *serious–carefree, subjective–analytic, energetic–relaxed, unassuming–self-asserting, lenient–firm, intense–calm, quiet–talkative, introverted–extroverted, steady–flexible, cautious–bold, cooperative–competitive*) to which participants responded by circling either one or the other of the two traits or by circling a third option—that is, *variable/depends on the situation*—indicating their reluctance to ascribe a stable personal attribute. The sample of university staff responded to similar questionnaire items, except that they were asked to assess only themselves at present or themselves 10 years ago.	p.199
2	They did not indicate their gender.[3] The procedure and measures resembled those of Study 1, except that participants were randomly assigned to provide assessments of themselves either 5 years into the future or of themselves in the present.	p.200
3	*Procedure and dependent measures.* Participants were randomly assigned to provide trait assessments of themselves in the present, 5 years ago, or 5 years from now. They provided their assessments on scales that were a variant of the Nisbett et al., (1973) scales used in Studies 1 and 2. The trait pairs were kept as faithful to their original meanings as possible but were modified so that each pair included one negatively toned trait (indicated below by an *N*) and one positively tone trait (indicated below by a *P*). The pairs thus became: *uptight (N)–easygoing (P), fickle (N)–reasonable (P), energetic (P)–lazy (N), shy (N)–self-assured (P), passive (N)–decisive (P), frenzied (N)–cool-headed (P), concise (P)–wordy (N), reclusive (N)–sociable (P), stubborn (N)–adaptable (P), timid (N)–brave (P), helpful (P)–selfish (N)*. (A research assistant unaware of our hypotheses and given a randomly ordered list of this entire set of traits rated each trait according to the predicted valence noted above.)	p.201

II.) Original Article Hypotheses and Corresponding Findings

Table S5

Pronin and Ross (2006): Summary of hypotheses in Studies 1-3

Study	Hypothesis	Description of hypothesis
1	Hypothesis 1a	Actor-observer asymmetry in self- vs other (friend) trait ascriptions: people make more trait ascriptions, as opposed to ascriptions to situational variability, when making assessments about other people compared to when they are making assessments about themselves.
	Hypothesis 1b	Temporal asymmetry resembles actor-observer asymmetry: trait attributions of one's past self resembles those made about other people more so than attributions made of one's present self. Meaning, people ascribe more traits, as opposed to ascriptions to situational variability, to their past-self and to other people compared to their present-self.
2	Hypothesis 2	Temporal asymmetry in trait self-ascriptions: people make more trait ascriptions, and less ascriptions to situational variability, to their future selves compared to their present selves.
3	Hypothesis 3a	Self-enhancement hypothesis of present self trait ascriptions: people ascribe more positive traits than negative traits to their present self.
	Hypothesis 3b	Self-enhancement hypothesis of temporal asymmetry in trait self-ascriptions: people ascribe more negative traits to their past-self, less trait attributions and more ascriptions to situational variability for their present-self, and more positive traits to their future-self.
	Hypothesis 3c	Temporal asymmetry: people ascribe more traits, regardless of whether they are negative or positive, to their past and future selves compared to their present-selves.

III.) Combining replication of Studies 1-3 into single data collection (random order)

Following our detailed analysis of Studies 1-3 (as presented in sections I-II), we decided to combine the designs of the three studies into a single data collection with randomization of the order of the combined studies (from now on referred to as "combined design").

This design allows for a direct replication of the three studies with important benefits compared to three separate replications. First, if one replication were to succeed whereas another would fail it would address any possible questions regarding the sample, especially regarding concerns of attentiveness and data quality, as the successful replication would clearly demonstrate that some effects are replicable with this sample. Second, the combined designed allows to address potential issues and allow us to test additional factors that would only be possible in a combined design, examining interactions and potential interplay between the studies.

In the following, we detail on our choice of how to integrate the different studies into a unified design while minimizing possible issues:

The combination of the different items in the dependent variables (DVs) of different studies into in a single DV question displaying items in randomized order builds on the original designs and allows stronger testing. We made sure to randomize the presentation of the adjective pairs. The DV was a combination of two sets of scales that were very similar in nature - both consisted of a list of adjectives used to describe an individual's personality. The main difference between them was that the pairs of adjectives in the second set of scales were oppositely valenced while the first set of scales were not, and so including them in a single randomized set actually ensures that participants are not blindly repeating answers. This also addresses concerns regarding order effects, allowing for testing of potential issues.

We combined the IVs from the different studies, given the between-subjects design where each participant was randomly allocated to only one of the six conditions, each participant did not experience the study any differently than those who completed the original studies other than having to rate a set of 22 scales instead of 11 scales (see point above regarding DVs).

As for the statistical approach, we followed and reproduced the data analysis method of the original studies as closely as possible. Although we added several tests (checks for outliers and assumption tests, non-parametric tests, the MANOVAs), this was done in addition to the original tests for the purposes of ensuring robustness. Moreover, when evaluating the replication results, we did so by using the results reproducing the original statistical methods that Pronin and Ross (2006) employed.

As such, we summarize these deviations as minor adjustments that should not have any major implications on the replication success and with important added benefits

Table S6

Pronin and Ross (2006): Summary of findings in Studies 1-3

Study	Factors	Effect	95% Confidence Interval	
			Lower	Upper
1	Temporal Distance x Social Distance (past-self vs present-self vs past-friend vs present-friend)	0.39	0.18	0.56
	Social Distance (self vs friend)	0.35	0.09	0.61
	Temporal Distance (past vs present)	0.32	0.00	0.62
2	Temporal Distance (present vs future)	0.51	0.17	0.84
3	Temporal Distance (past vs present vs future)	0.46	0.16	0.75
	Temporal Distance (present vs future)	0.54	0.27	0.77
	Temporal Distance (past vs present)	0.59	0.28	0.89
	Attribution of Positive vs Negative Traits (past vs present vs future)	0.16	0.00	0.36
	Attribution of Positive vs Negative Traits (positive vs negative)	0.77	0.29	1.25

Note. Effect = Cohen's *f*. The effect sizes presented were calculated based on the reported test statistics from the original article. Please refer to the section on "effect size calculations" in the supplementary for more details regarding those calculations.

IV.) Original Article Results

A.) Sample size before and after exclusions

No exclusions were reported by the authors in studies 1-3.

- Study 1: *N* = 170 (two samples: 123 undergraduate students and 47 university staff)
- Study 2: *N* = 40
- Study 3: *N* = 75

B.) Included sample description

Please refer to table S7 for details about the sample demographics for studies 1-3, and table S8 for the sample description included in the original article.

Table S7

Sample descriptions of studies 1-3 of Pronin & Ross (2006)

Descriptor	Study 1		Study 2	Study 3
	Undergraduate sample	University staff sample		
Age*	19	44	not reported	not reported
Gender	66% female, 34% male	65 females, 3 males	64% female, 36% male	39 females, 35 males, 1 not reported
Location	Online. Participants responded to an email recruitment and completed a web-based questionnaire.			
Sample type	Undergraduate university students taking a psychology course	University academic staff	Undergraduate university students taking a psychology course	Undergraduate university students

*Note: median age is reported, mean and standard deviation was not reported in the original article.

Table S8

Participants of studies 1-3 of Pronin & Ross (2006) as described in the original article.

Study	Participants as described in Pronin & Ross (2006)	Page no.
1	*Participants.* Two respondent groups were used. One was a group of college undergraduates who participated to fulfill a course requirement (n = 123; median age = 19). The other was a group of university staff working in various academic departments who responded to an e-mail asking them to volunteer their help in completing a Web-based questionnaire (n = 47; median age = 44). Undergraduates did not indicate their gender;[1] the staff sample contained 65 women and 3 men. [1] The sample contained 66% women and 34% men.	p.199
2	A total of 40 undergraduates participated in exchange for course credit. They did not indicate their gender.[3] The procedure and measures resembled those of Study 1, except that participants were randomly assigned to provide assessments of themselves either 5 years into the future or of themselves in the present. [3] The sample contained 64% women and 36% men.	p.200
3	*Participants.* A total of 75 undergraduates (39 women, 35 men, and 1 participant who did not indicate gender[4]) participated in exchange for their choice of candy bars.	p.201

C.) Descriptive Statistics

Please refer to table S9 for an overview of the descriptive statistics of studies 1-3, table 8 for the bar charts visualizing the descriptive statistics taken from the original article, and table S10 to see the results as described in the original article by Pronin & Ross (2006).

Number of participants. Since design is between-subjects and randomized, it is assumed that participants are equally divided among the conditions. Since the samples are odd-numbered and conditions are even, one condition will have an extra participant at random.

- Study 1:
  - student sample : 123/4 conditions = 30 participants per condition
  - staff sample : 47/2 conditions = 23 participants per condition
- Study 2:
  - 40/2 conditions = 20 participants per condition
- Study 3:
  - 75/2 conditions = 37 participants per condition

Table S9

Descriptive statistics of studies 1-3 of Pronin & Ross (2006)

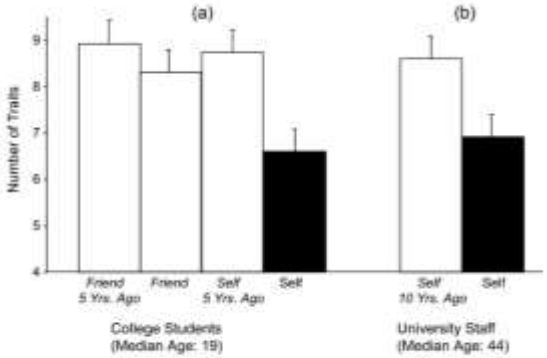
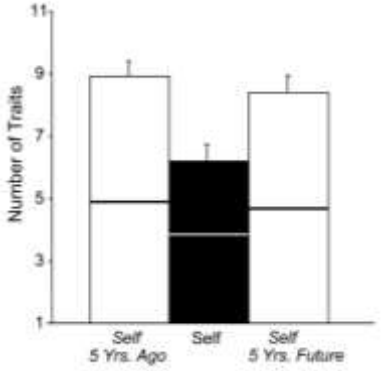
Study	Condition	DV1: No. of trait attributions		DV2: No. of positive trait attributions
		University staff sample (Study 1)	Undergraduate sample	
1	**Past-self**	*M* = 8.61	*M* = 8.75	-
	Present-self	*M* = 6.92	*M* = 6.61	-
	Past-friend	-	*	-
	Present-friend	-	*M* = 8.32	-
2	**Present-self**	-	*M* = 5.64	-
	Future-self	-	*M* = 8.17	-
3	**Past-self**	-	*M* = 8.92	*M* = 6.05** (55% positive)
	Present-self	-	*M* = 6.20	*M* = 6.60** (60% positive)
	Future-self	-	*M* = 8.40	*M* = 6.27** (57% positive)

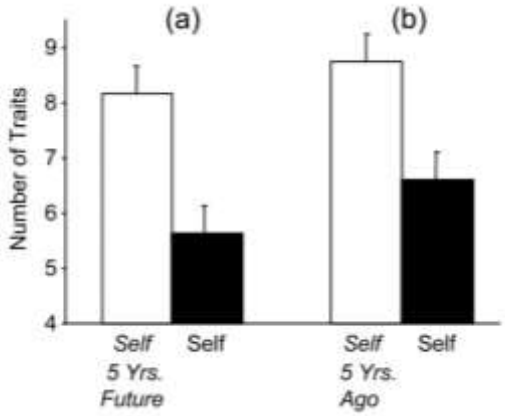
*Note : mean of past-friend condition is not reported in text but is displayed in the bar graph in table S8.

**Means of DV2 were calculated by multiplying the mean percentage of positive traits attributed by 11.

Table S10

Screenshots of bar charts of descriptive statistics taken from the original article

Study 1	**Study 2**
 *Figure 1.* Number of traits attributed (Study 1). Panels show (a) college students considering themselves (and a friend) in the past versus in the present and (b) university employees considering themselves in the past versus in the present. Error bars indicate 1 standard error above the mean.	 *Figure 2.* Number of traits attributed (Studies 1 and 2). Panels show college students assessing themselves (a) in the future versus in the present (Study 2) and (b) in the past versus in the present (Study 1). Error bars indicate 1 standard error above the mean.
(Source : p.199)	(Source : p.200)

Study 3
 *Figure 3.* Number of positive and negative traits attributed to the self in the past, present, and future (Study 3). The number of positive traits is indicated by the space below the horizontal line on each bar, and the number of negative traits is indicated by the space above that line. Error bars indicate 1 standard error above the mean.
(Source : p.201)

D.) Statistical test results

Please refer to table S11 to see the statistical test results for studies 1-3, and table S12 for screenshots of the results section of studies 1-3 from the original article. The below tests were analyzed by the original authors using between-subjects ANOVAs, and an additional independent t-test for study 3.

Table S11

Reported test statistics across studies 1-3 of Pronin & Ross (2006)

Study	Conditions	Target effect	*df*	*F*-statistic/ t-statistic	*p*
1	student sample: All conditions	Psychological distance (temporal vs social)	3	*F* = 6.23	.0006
	student sample: past-self vs present-self	Temporal distance (past vs present)	1	*F* = 11.63	.001
	student sample : present-self vs present-friend	Social distance (self vs friend)	1	*F* = 7.42	.008
2	staff sample : past-self vs present-self	Temporal distance (past vs present)	1	*F* = 4.74	.03
	present-self vs future-self	Temporal distance (present vs future)	1	*F* = 9.97	.003
	All conditions	Temporal distance (past vs present vs future)	2	*F* = 10.55	< .0001
	present-self vs future-self	Temporal distance (present vs future)	1	*F* = 10.37	.002
	past-self vs present-self	Temporal distance (past vs present)	1	*F* = 16.82	.0002
3	past-self vs future-self	Temporal distance (past vs future)	exact statistics not provided: "Past and future attributions did not differ from each other, *F* < 1." (p. 201)		
	All conditions	Attribution of positive vs negative traits (past vs present vs future)	2	*F* = 0.97	*ns*
	present-self	Attribution of positive vs negative traits (positive vs negative)	73	*t* = 3.30	.0002

Table S12

Results of studies 1-3 of Pronin & Ross (2006) as described in the original article.

Study	Results as described in Pronin & Ross (2006)	Page no.
1	Each participant's responses were first converted to a single score corresponding to the number of trait ascriptions that were made (rather than ascriptions to situational variability). *Undergraduate sample.* Undergraduates' trait ascriptions revealed an overall tendency to differ depending on the social and temporal distance of their judgment, $F(3, 120) = 6.23$, $p = .0006$. Before examining ascriptions to the present versus the past self, we tested whether participants did indeed show an actor–observer difference in their ascriptions. Consistent with earlier research, participants rated themselves as currently exhibiting fewer traits ($M = 6.61$), and, thus, as exhibiting more cross-situational variability, than their friends currently exhibited ($M = 8.32$), $F(1, 61) = 7.42$, $p = .008$. More relevant to our present concerns, we also found a similar asymmetry in assessments made about the present self ($M = 6.61$) versus the past self ($M = 8.75$), $F(1, 62) = 11.63$, $p = .001$. As Figure 1a illustrates, participants appeared to see their present selves as uniquely free of stable, cross-situational dispositions. This observation was supported by Tukey's method of individual pairwise comparisons. Participants' views of their present selves differed from their views of themselves in the past, their friend in the past, and their friend in the present (*p*s ranging from 8.03 to 10.90, *p*s $< .05$). No other differences were apparent. *University staff sample.* The data presented in Figure 1b suggest that our findings do not reflect the unique perspective of college students looking back on their less worldly, less flexible 14-year-old selves. Our sample of university staff, whose average age was 44, showed a similar asymmetry in indicating the number of traits they were willing to ascribe to their present selves ($M = 6.92$) versus their selves of 10 years ago ($M = 8.61$), $F(1, 46) = 4.74$, $p = .03$. Thus, just as our 19-year-old participants reported that when they were 14 years old, they had nine traits but that now at 19 they had seven traits, our 44-year-olds reported that at 34 they had nine traits and it was only now, at age 44, that they had seven traits. In short, it was the contrast of present and past selves, not the contrast of two particular ages, that produced the relevant asymmetry.	p.199
2	Consistent with our hypothesis, participants considering themselves 5 years into the future claimed that they would manifest more of the 11 possible traits ($M = 8.17$) than did participants considering their current selves ($M = 5.64$), $F(1, 39) = 9.97$, $p = .003$. As can be seen in Figure 2, our participants' perceptions of their future versus present selves closely resembled our Study 1 participants' perceptions of their past versus present selves.	p.200
3	Consistent with the results of Studies 1 and 2, participants considering themselves in the present ascribed fewer traits to themselves ($M = 6.20$) than did participants considering themselves 5 years into the future ($M = 8.40$) or 5 years into the past ($M = 8.92$), $F(2, 73) = 10.55$, $p < .0001$. These differences were significant for both present versus future attributions, $F(1, 49) = 10.37$, $p = .002$, and present versus past attributions, $F(1, 49) = 16.82$, $p = .0002$. Past and future attributions did not differ from each other, $F < 1$. *Attribution of positive and negative traits.* We next examined whether participants showed any temporal differences in their tendency to ascribe positive versus negative traits (see Figure 3). To do this, we examined participants' ratios of the number of positive traits they selected to total traits selected. Participants generally attributed a favorable ratio of positive-to-negative traits to themselves, $t(73) = 3.30$, $p = .0002$, but this tendency did not differ for attributions about present self (60% positive/40% negative), future self (57%/43%), and past self (55%/45%), $F(2, 73) = .97$, *ns*.	p.201

V.) Effect size calculations of the original study effects

Effect sizes were calculated in R studio (R Core Team, 2013) using the *F_to_f()* and *t_to_f()* functions from the *effectsize* package (Ben-Shachar, Makowski & Lüdecke, 2020) based on the test statistics reported in the original article. Table S13 displays : 1.) the reported test statistics used to calculate effect sizes, 2.) the R code used to calculate each effect size and 95% C.I., and 3.) the corresponding calculated effect sizes and 95% C.I . Please refer to the OSF to find the original R-script (filename: Effect size and power analysis.Rmd).

Table S13

Calculation and findings of original study effects

Study	Target effect	Reported test statistic	R- code used	Calculated effect sizes and 95% CI
1	Student sample: Psychological distance (temporal vs social)	*F* (3, 120) = 6.23	F_to_f(f=6.23, df=3, df_error=120, ci = 0.95)	Cohen's *f* = 0.39, 95% C.I. [0.18, 0.56]
	Student sample: Temporal distance (past vs present)	*F* (1, 62) = 11.63	F_to_f(f=11.63, df=1, df_error=62, ci = 0.95)	Cohen's *f* = 0.43, 95% C.I. [0.17, 0.69]
	Staff sample: Temporal distance (past vs present)	*F* (1, 46) = 4.74	F_to_f(f=4.74, df=1, df_error=46, ci = 0.95)	Cohen's *f* = 0.32, 95% C.I. [0.00, 0.62]
	Student sample: Social distance (self vs friend)	*F* (1, 61)= 7.42	F_to_f(f=7.42, df=1, df_error=61, ci = 0.95)	Cohen's *f* = 0.35, 95% C.I. [0.09, 0.61]
2	Temporal distance (present vs future)	*F* (1, 39) = 9.97	F_to_f(f=9.97, df=1, df_error=39, ci = 0.95)	Cohen's *f* = 0.51, 95% C.I. [0.17, 0.84]
3	Temporal distance (past vs present vs future)	*F* (2, 73) = 10.55	F_to_f(f=10.55, df=2, df_error=73, ci = 0.95)	Cohen's *f* = 0.54, 95% C.I. [0.27, 0.77]
	Temporal distance (present vs future)	*F* (1, 49) = 10.37	F_to_f(f=10.37, df=1, df_error=49, ci = 0.95)	Cohen's *f* = 0.46, 95% C.I. [0.16, 0.75]
	Temporal distance (past vs present)	*F* (1, 49) = 16.82	F_to_f(f=16.82, df=1, df_error=49, ci = 0.95)	Cohen's *f* = 0.59, 95% C.I. [0.28, 0.89]
	Attribution of positive vs negative traits (past vs present vs future)	*F* (2, 73) = 0.97	F_to_f(f=0.97, df=2, df_error=73, ci = 0.95)	Cohen's *f* = 0.16, 95% C.I. [0.00, 0.36]
	Attribution of positive vs negative traits (positive vs negative)	*t* (73) = 3.30	t_to_d(t=3.30, df_error=73, ci = 0.95)	Cohen's *d* = 0.77, 95% C.I. [0.29, 1.25]

VI.) Power analysis of original study effect to assess required sample for replication

We concluded that the minimum required sample size for a power of 0.95 and alpha of 0.05 is 362 participants. Table S12 displays the test statistics, effect sizes, and R-code used to conduct the power analyses and calculate the required sample size for the current study.

We calculated the 60% C.I. using the F_to_f() function mentioned in the previous section with the confidence interval set to 0.60 instead of 0.95, and then we conducted our power analysis using the pwr.anova.test() function from the *pwr* package (Champely et al., 2018) in R (R Core Team, 2013). Please refer to the OSF to find the original R-script (filename: Effect size and power analysis.Rmd).

Table S14

Calculating lower-bound effect size and power analysis based on smallest calculated effect sizes from table 11

Study	Target effect	Reported test statistic	R-code used (60% C.I. effect size)	Calculated 60% Confidence Interval	R-code used (power analysis)	Required sample size (per group)	Total required sample size
1	Staff sample: Temporal distance (past vs present)	*F* (1, 46) = 4.74	F_to_f(f=4.74, df=1, df_error = 46, ci = 0.60)	0.19, 0.45	pwr.anova.test(k = 2, n = NULL, f = 0.19, sig.level = 0.05, power = 0.95)	181	181*2= 362
2	Temporal distance (present vs future)	*F* (1, 39) = 9.97	F_to_f(f=9.97, df=1, df_error = 39, ci = 0.60)	0.36, 0.65	pwr.anova.test(k = 2, n = NULL, f = 0.36, sig.level = 0.05, power = 0.95)	51	51*2= 102
3	Temporal distance (present vs future)	*F* (1, 49) = 10.37	F_to_f(f=10.37, df=1, df_error = 49, ci = 0.60)	0.33, 0.58	pwr.anova.test(k = 3, n = NULL, f = 0.16, sig.level = 0.05, power = 0.99)	61	61*2= 122

* Note: sample size rounded to the nearest whole number

VII.) Sensitivity analysis

We used the pwr.t.test() from the *pwr* package in R to conduct post-hoc sensitivity analysis. Setting *n* as 291, the replication for Study 1 was sensitive to detect an effect of *d* = 0.39 with 95% power and effect of 0.29 with 80% power (one-sided).

```
> # for 95%
> (estimate <- pwr.t.test(n = 145,
+                         sig.level = 0.05,
+                         power = 0.95,
+                         type = "two.sample",
+                         alternative = "greater"))

     Two-sample t test power calculation

              n = 145
              d = 0.3872707
      sig.level = 0.05
          power = 0.95
    alternative = greater

NOTE: n is number in *each* group
> # for 80%
> (estimate <- pwr.t.test(n = 145,
+                         sig.level = 0.05,
+                         power = 0.80,
+                         type = "two.sample",
+                         alternative = "greater"))

     Two-sample t test power calculation

              n = 145
              d = 0.2927114
      sig.level = 0.05
          power = 0.8
    alternative = greater

NOTE: n is number in *each* group
```

Materials and scales used in the replication + extension experiment

Procedure

Participants were recruited through the Amazon MTurk platform and were asked to complete a Qualtrics survey in exchange for a small monetary compensation. The procedure of the survey and the differences in instructions between conditions is presented in Table S15.

Instructions and experimental material

Please refer to the OSF to find the original Qualtrics used in the present study. Table 14 presents a summary of the survey flow, along with the instructions given to the participants in each experimental condition. Please refer to table S17 to view the set of scales participants had to complete in all experimental conditions.

Table of design : 2 (temporal distance) x 2 (social distance) between-subjects experimental design

Table S15

Table of design of the current replication and extension

Psychological Distance IV1: Temporal distance (past vs present vs future) IV2: Social distance (self vs friend)	**Manipulation example:** In this part of the study, you will be asked to provide assessments about [insert IV2 manipulation]'s personality on a set of 22 scales. Each scale consists of three descriptors (e.g. serious, carefree, variable/depends on the situation), and you will be asked to choose one of the three which you think best described [IV2 manipulation] [insert IV1 manipulation]		
	IV1: Past (between) [five years ago.]	**IV1 : Present (between)** [right now.]	**IV2: Future (between)** [five years from now.]
IV2: Self (between) [yourself]	**DV1: Attributional style (actor-like : trait ascriptions vs observer-like : ascriptions to situational variability)** Specific DV item: the number of trait ascriptions (e.g. serious, subjective) made (instead of ascribing to situational variability - choosing "variable/depends on the situation").	**Set of 22 scales : adapted from Nisbett et al. (1973)** 1. serious– carefree—variable/depends on the situation 2. subjective–analytic—variable/depends on the situation 3. energetic–relaxed—variable/depends on the situation 4. unassuming–self-asserting—variable/depends on the situation 5. lenient–firm—variable/depends on the situation 6. intense– calm—variable/depends on the situation 7. quiet–talkative—variable/depends on the situation 8. introverted– extroverted—variable/depends on the situation 9. steady–flexible—variable/depends on the situation 10. cautious– bold—variable/depends on the situation 11. cooperative– competitive —variable/depends on the situation 12. uptight (N)–easygoing (P)–variable/depends on the situation 13. fickle (N)– reasonable (P)–variable/depends on the situation 14. energetic (P)–lazy (N)–variable/depends on the situation 15. shy (N)–self-assured (P)–variable/depends on the situation 16. passive (N)–decisive (P)–variable/depends on the situation 17. frenzied (N)–cool-headed (P)–variable/depends on the situation 18. concise (P)–wordy (N)–variable/depends on the situation 19. reclusive (N)–sociable (P)–variable/depends on the situation 20. stubborn (N)–adaptable (P)–variable/depends on the situation 21. timid (N)–brave (P)–variable/depends on the situation 22. helpful (P)–selfish (N)–variable/depends on the situation	
IV2 : Friend (between) [the initials of a friend that you have known for at least five years]	**DV2: Attribution of positive and negative traits** Specific DV item: the number of positive traits (denoted by "(P)", as opposed to "(N)" for negative traits) to total traits selected.		

Table S16

Summary of Qualtrics survey

Step	Survey Flow			
1.)	**Consent form** : Participants read the consent form which explained the purpose of the study, procedure, potential risks, potential benefits, compensation, confidentiality, participation, and withdrawal, and how to get in touch if they have questions or concerns.			
2.)	**Survey introduction** : Participants read the outline of the overall study, explaining its content and length, along with some verifications to ensure that they are paying attention.			
3.)	**Specific condition instructions and comprehension checks:** Participants are randomly assigned to one of the 6 possible experimental conditions. Prior to seeing the manipulation, they viewed the instructions for the specific condition they have been assigned to, and then were asked to answer comprehension checks that assessed whether they knew who they were supposed to rate (IV2) and at what time point (IV1).			
	Part 1 : Replication of studies 1-3 from Pronin & Ross (2006) Please reference table 15 to see the DV measure used in all conditions.			
4.)	IV1: Temporal Distance IV2: Social Distance	**IV1: Past**	**IV1: Present**	**IV1: Future**
	IV2: Self	**Exp .Condition 1 (Past-Self)** Participants were asked to give assessments about themselves 5 years ago on a set of 22 scales.	**Exp. Condition 2 (Present-Self)** Participants were asked to give assessments about themselves right now on a set of 22 scales.	**Exp. Condition 3 (Future-Self)** Participants were asked to give assessments about themselves five years from now on a set of 22 scales.
	IV2: Friend	**Exp .Condition 4 (Past-Friend)** Participants were asked to think of a friend they have known for five years, five years ago, whom they are still friends with, and write their initials in a text box. This was piped to the next instruction, which asked them to give assessments about this	**Exp .Condition 5 (Present-Friend)** Participants were asked to think of a friend they have known for five years, and write their initials in a text box. This was piped to the next instruction, which asked them to give assessments about this friend right now on a set of 22 scales.	**Exp .Condition 6 (Future-Friend)** Participants were asked to think of a friend they have known for five years, and write their initials in a text box. This was piped to the next instruction, which asked them to give assessments about this

	friend 5 years ago on a set of 22 scales.	friend five years from now on a set of 22 scales.

Part 2 : Manipulation check of measure items from study 3 of Pronin & Ross (2006)

5.) All participants were asked to rate 22 traits (from study 3) in terms of their valence (how desirable/undesirable they are). Please reference table 16 to see the scale used.

6.) **Funneling section:** Participants were asked the following questions : "What do you think the purpose of the last part was?", "Have you ever seen the materials used in this study or similar before? If yes - please indicate where.", " Did you spot any errors? Anything missing or wrong? Something we should pay attention to in next runs? (Briefly, up to one sentence, write "none" if not relevant)." Afterwards, participants were asked to fill out their demographic information, thanked for their participation, and debriefed.

Scales used in the experiments

Table S17

Set of scales presented to participants in the current replication and extension.

DV and details of the measurement	Scales
For the present study, we combined the measures used in study 1, 2, and 3 into one set of scales instead of two. We then randomize the order of the items in the experiment. **<u>DV1: Attributional style</u>** For each participant, their responses are summarized into one score denoting the amount of trait attributions they made. ● actor-like : greater number of trait ascriptions, compared to situational variability ● observer-like : greater number of ascriptions to situational variability than trait ascriptions. **<u>DV2: Attribution of positive and negative traits</u>** Measures the ratio of positive (P) to negative (N) traits participants ascribed to themselves in the past/present/future (depending on the condition they were assigned to). ● self-enhancement hypothesis: participants describe their past self with more negative traits, their present self as more variable, and their future self with more positive traits.	Adapted from 11 scales measuring observer-actor difference in trait ascriptions from Nisbett et al. (1973) 1. serious– carefree—variable/depends on the situation 2. subjective–analytic—variable/depends on the situation 3. energetic–relaxed—variable/depends on the situation 4. unassuming–self-asserting—variable/depends on the situation 5. lenient–firm—variable/depends on the situation 6. intense– calm—variable/depends on the situation 7. quiet–talkative—variable/depends on the situation 8. introverted– extroverted—variable/depends on the situation 9. steady–flexible—variable/depends on the situation 10. cautious– bold—variable/depends on the situation 11. cooperative– competitive —variable/depends on the situation 12. uptight (N)–easygoing (P)–variable/depends on the situation 13. fickle (N)– reasonable (P)–variable/depends on the situation 14. energetic (P)–lazy (N)–variable/depends on the situation 15. shy (N)–self-assured (P)–variable/depends on the situation 16. passive (N)–decisive (P)–variable/depends on the situation 17. frenzied (N)–cool-headed (P)–variable/depends on the situation 18. concise (P)–wordy (N)–variable/depends on the situation 19. reclusive (N)–sociable (P)–variable/depends on the situation 20. stubborn (N)–adaptable (P)–variable/depends on the situation 21. timid (N)–brave (P)–variable/depends on the situation 22. helpful (P)–selfish (N)–variable/depends on the situation

Table S18

Material for manipulation check of DV2 : trait rating of 22 opposite valence traits from study 3 of Pronin & Ross (2006)

Manipulation check	Scales
	Contains all the trait terms from the scales used in study 3 of Pronin & Ross (2006). 1. uptight 12. cool-headed 2. easygoing 13. concise 3. fickle 14. wordy 4. reasonable 15. reclusive 5. energetic 16. sociable 6. lazy 17. stubborn 7. shy 18. adaptable 8. self-assured 19. timid 9. passive 20. brave 10. decisive 21. helpful 11. frenzied 22. selfish
<u>Manipulation check for DV2: Attribution of positive and negative traits</u> This measures the valence (how positive or negative) each trait is perceived to be by the participants, and is used to confirm whether the ascribed valences of the trait terms from study 2 are accurate and applicable to the MTurk sample. This measure was presented to the participants at the end of the survey.	For each trait, participants are asked to rate the desirability of the trait term on a scale of -3 (very negative) to 3 (very positive) -3 (very negative) | -2 | -1 | 0 | 1 | 2 | 3 (very positive)

Exclusion criteria

Generalized exclusion criteria

The default generalized exclusion criteria we use in our pre-registration is the following:

"We will focus on our analyses on the full sample. However, as a supplementary analysis and to examine any potential issues, we will also determine further findings reports with exclusions. In any case, we will report exclusions in detail with results for the full sample and results following exclusions (in either the manuscript or the supplementary).

General criteria:

1. Participants indicating a low proficiency of English (self-report < 5, on a 1-7 scale)
2. Participants who self-report not being serious about filling in the survey (self-report < 4, on a 1-5 scale).
3. Participants who correctly guessed the hypothesis of this study in the funneling section.
4. Participants who have already seen or done the survey before.
5. Participants who failed to complete the survey. (duration = 0, leave question blank)
6. (When target sample is MTurk:) Participants not from United States.

Comparisons and deviations

Original versus replication

Item	Original	Replication	Reason for change
Study design	3 separate designs: ● Study 1: 2 (past vs present) x 2 (self vs friend) between-subjects design ● Study 2: present-self vs future self, between-subjects design ● Study 3 : mixed design. Past vs present vs future self between subjects design, and negative vs positive trait attributions as within subjects.	3 (temporal distance : past vs present vs future) x 2 (social distance: self vs friend) between-subjects design.	Combined studies 1-3 into one study due to similar experimental design that do not conflict with each other
Conditions	Participants randomly assigned to one of… ● four experimental conditions (past-friend vs present-friend vs past-self vs future-self) in study 1	Participants randomly assigned to one of six experimental conditions: ● 6 between-subject experimental conditions in total : past-self vs present-self vs future-self vs past-friend vs present-friend vs future-friend ● (Added future-friend condition as an extension)	Same as above
Procedure	● Participants were told to answer the scales according to the condition they were randomly assigned to. ● (RA asked to verify manipulation for IV2)	Same instructions as the original with a few additions: ● Added comprehension check throughout the survey ● Added manipulation check for IV2 (positive vs negative traits) at the end of the survey	To account for the MTurk sample, we had to add comprehension checks to make sure participants were paying attention. Manipulation check was also added since in the original study they had an RA verify the manipulation, but it may not be applicable for MTurk sample.
Measures	Online web survey: ● Study 1 and study 2 used the same scales. Study 3 used a modified scale with opposite valenced traits. ● 1 measure has 11 scales ● Items are in fixed order	Online Qualtrics survey: ● Combined scales from study 1 and 2 and study 3, into one set of scales with the order of the items randomized. ● 1 measure has 22 scales	Since we combined the experimental design of studies 1-3, we combined the two measures in one, in randomized order, as separating them would have statistical implications.
Participant population	Undergraduate university students and university staff (only study 1) recruited through university database email	Recruited through the online platform Amazon MTurk	To make the results more generalizable beyond university populations. Since MTurk samples are more diverse age wise, we did not have to separate them into two sample groups (university staff and university students)

Pre-exclusions versus post-exclusions

Overview of pre-exclusions and post-exclusions

	Before exclusion	Fulfilled any of the exclusion criteria	After exclusion
Total number of cases	911	33	878

Summary of exclusions

Exclusion Criteria	Cases fulfilling the exclusion criteria*
Exclusion Criteria 1a Participants indicating low proficiency of English.	5
Exclusion Criteria 1b Participants who self-report not being serious about filling in the survey.	11
Exclusion Criteria 1c Participants who correctly guessed any one of the hypotheses of this study in the funneling section.	0
Exclusion Criteria 1d Participants who have already seen the materials in the survey before	10
Exclusion Criteria 1e Participants who failed to complete the survey.	0
Exclusion Criteria 1f Participants not from the United States.	10

**Note* : some cases overlap in regard to fulfilling the exclusion criteria.

Table S19

Summary of pre-exclusion versus post-exclusion main results for the replication condition

Study	Hypothesis		Target effect	Pre-exclusion (Cohen's *f)*	Post-exclusion (Cohen's *f)*
	No.	Description			
1	1a	Actor-observer asymmetry in self- vs other (friend) trait ascriptions	Social distance (self vs friend)	0.10, 95% CI [0.04, 0.17]	0.10, 95% CI [0.04, 0.17]
	1b	Temporal asymmetry resembles actor-observer asymmetry	Temporal distance (past vs present)	-0.02, 95% C.I. [-0.02, -0.25]	-0.01, 95% C.I. [-0.21, 0.19]
2	2	Temporal asymmetry in trait self-ascriptions	Temporal distance (present vs future)	0.03, 95% C.I. [-0.20, 0.25]	0.03, 95% C.I. [-0.20, 0.25]
3	3a	Self-enhancement hypothesis of temporal asymmetry in trait self-ascriptions	Ratio of positive-to-total trait ascriptions (present self)	0.88, 95% CI [0.50, 1.26]	0.87, 95% CI [0.50, 1.24]
	3b		Ratio of positive-to-total trait ascriptions (past vs present vs future self)	0.33, 95% CI [0.22, 0.42]	0.33, 95% CI [0.22, 0.42]
	3c	Temporal asymmetry	Temporal distance (past vs present vs future)	0.02, 95% CI [0.00, 0.06]	0.01, 95% CI [0.00, 0.05]

Note. We did not find any significant differences in the main results between the pre-exclusion and post-exclusion analysis.

Pre-registration plan versus final report

Components in your preregistration (e.g., stopping rule, analyses, hypotheses, exclusion rules)	Location of 1) preregistered decision/ plan and 2) rational for decision/ plan [Location / link]	Were there deviations? What type? [no / minor / major] *	If yes - describe details of deviation(s) [brief description / location / link]	Rationale for deviation [brief description / location / link]	How might the results be different if you had/had not deviated [brief description / location / link]	Date/time of decision for deviation + stage
Study design		no				
Measured variables		no				
Exclusion criteria		no				
IV		no				
DV		minor	For the self-enhancement hypothesis, DV was changed from positive traits to ratio of positive-to-negative traits.	Follows the original article analysis more closely for the self-enhancement hypothesis.	Small difference	
Data analysis		minor	Calculation of the ratio of positive over total traits ascribed for the self-enhancement hypothesis. In the pre-registration dummy data analysis, DV was calculated using the number of positive traits but not the ratio of positive-to-total traits.			
Data analysis		minor	Following the MANOVA, added in 2x3 ANOVA for each DV in addition to the already existing ANOVA for each IV.	To further explain the variables underlying the significant MANOVA term	Full reason underlying significant MANOVA term would be unclear	After first round of peer review (06/2021)

Evaluation criteria for replication findings

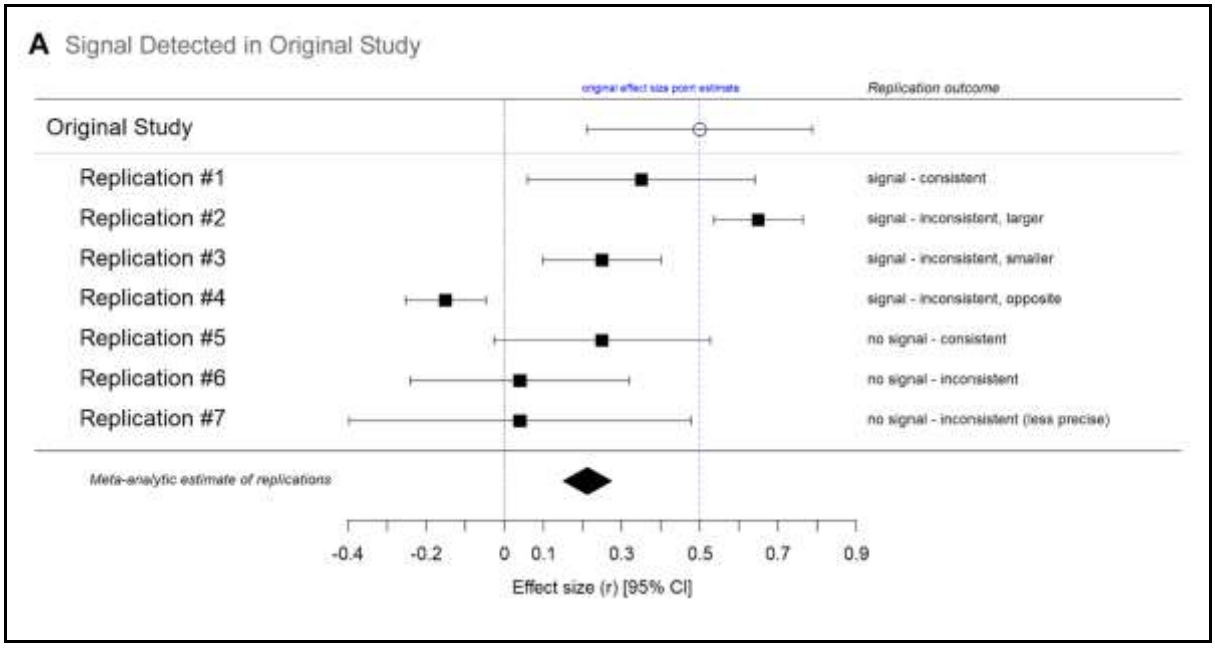


Figure S2. Criteria for evaluation of replications by (LeBel et al., 2019). A simplified replication taxonomy for comparing replication effects to target article original findings.

Target similarity	Highly similar				Highly dissimilar
Category	Direct replication			Conceptual replication	
Design facet	Exact replication	Very close replication	Close replication	Far replication	Very far replication
IV operationalization	Same/similar	Same/similar	Same/similar	Different	
DV operationalization	Same/similar	Same/similar	Same/similar	Different	
IV stimuli	Same/similar	Same/similar	Different		
DV stimuli	Same/similar	Same/similar	Different		
Procedural details	Same/similar	Different			
Physical setting	Same/similar	Different			
Contextual variables	Different				

Figure S3. Criteria for evaluation of replications by LeBel et al. (2018).
A classification of relative methodological similarity of a replication study to an original study. “Same” (“different”) indicates the design facet in question is the same (different) compared to an original study. IV = independent variable. DV = dependent variable. “Everything controllable” indicates design facets over which a researcher has control. Procedural details involve minor experimental particulars (e.g., task instruction wording, font, font size, etc.).
"Similar" category was added to the Lebel et al. (2018) typology to refer to minor deviations or extensions aimed to adjust the study to the target sample that are not expected to have major implications on replication success. See Olsson-Collentine, van Assen, and Wicherts (2020) on meta analysis showing minor to no expected impact due to variations in sample population or setting.

Pre-registered plan for detecting and managing outliers and assumptions

Detecting and managing outliers

Prior to analysis, we screened the data for outliers using robust statistical methods of outlier identification with the *Routliers* package (Delacre & Klein, 2019). Following the recommendations of Leys et al. (2019), we detected univariate outliers using Median Absolute Deviation (MAD) with the *outliers_mad()* function, and multivariate outliers using the Minimum Covariance Determinant (MCD) with the *outliers_mcd()* function. For both functions, we set the threshold to the default number (threshold = 3 and h =.75 respectively). These methods are robust because they are based on the median, instead of usual practices that base outlier detection on the mean and standard deviation (e.g., Mahalanobis, 1930; Tabachnick & Fidell, 2013), which can be problematic because these values themselves are influenced by outliers (Leys et al., 2019). If we detected a significant number of outliers (>5%), we followed the recommendations from Mair and Wilcox (2019) and used robust statistical tests instead using the *WRS2* package (Mair & Wilcox, 2020) to account for the outliers. If not, we kept them in the analysis given that the data is normally distributed, and outliers are less than 5%.

Assumption checks of statistical tests

In their original article, Pronin and Ross (2006) did not report whether their data met the assumptions necessary for running their statistical tests. Since the data is in the form of counts (number of trait ascriptions), it is likely to violate the assumption of normality. Count data cannot be less than zero, and as such are likely to be highly skewed. Since violating important assumptions may invalidate the use of parametric tests and lead to incorrect inferences (Olsen, 2003), we first screened the data using various functions and

packages in R before proceeding with data analysis. We also provide additional alternative tests to run in case any of the critical test assumptions are not met. In any case, if any are violated, we will first report the original analysis plan following Pronin and Ross (2006), in addition to the alternative analyses in the supplementary to see whether the results will still hold.

Following the procedure described by Schumacker (2015) regarding assumption checks for multivariate statistical tests, we first checked the data for normality, skewness, and kurtosis using the function *myshapiro.test()* from the *mvnorm* package and the function *normality.test1()* from the *normwhn* package respectively. Then, we checked for the equality of variance- covariance matrices. First, we used the *cov()* and *det()* functions to calculate the variance- covariance matrix and determinant of matrix for each of the six groups, and then used the *boxM()* function from the *heplots* package to test the equality of covariance matrices between groups.

If all assumptions were met, we proceeded to run the analyses described below using the usual parametric tests in R following Pronin and Ross (2006). However, if only the assumption of normality was not met, we will run the statistical tests following a poisson distribution, which is more suitable for count data (Gardner, Mulvey, & Shaw, 1995). Following recommendations by Crawley (2005), we will use the functions *glm()* instead of *aov()* or *t.test()*, and specify *family=* as *poisson* to run the tests following a poisson distribution.

If, additionally, we found that assumptions of skewness or kurtosis were not met, we instead used robust statistical methods with the *WRS2* package (Mair & Wilcox, 2020) following the recommendations of Field and Wilcox (2016). Specifically, we used the *t1way()*

and *mcbpp()* function instead of *aov()* and *pairwise.t.test()* for conducting ANOVA, and *yuenbt()* instead of *t.test()* for conducting t-test. Unlike parametric tests used in Pronin and Ross (2006), these alternative tests are robust to violations of the aforementioned assumptions.

Additional analyses and results

Statistical assumptions and normality tests

Following the pre-registered data analysis plan, we first screened the data for both multivariate and univariate outliers in R. Since outliers were less than 5% (see Table S17), we retained them in the analysis. Afterwards, we checked whether the data meets the assumptions required for both multivariate and univariate statistical analysis. Please refer to Table S20 for a summary of the findings of these assumption checks.

Table S20

Summary of outlier detection and assumptions check

Type		Variable	Value	*p*	Met
Outliers	Univariate (MAD)	Total traits	0 outliers	Not applicable	Yes
		Total positive traits	8 outliers	Not applicable	Yes
	Multivariate (MCD)	Total traits and positive traits	23 outliers	Not applicable	Yes
Assumptions	Normality	Total traits	0.93	<.001	No
		Total positive traits	0.97	<.001	No
		Total negative traits	0.87	<.001	No
	Skew	Total traits	-0.7	<.001	No
		Total positive traits	-0.26	.023	No
		Total negative traits	1.08	<.001	No
	Kurtosis	Total traits	3.70	<.001	No
		Total positive traits	2.41	.076	Yes

	Total negative traits	3.81	<.001	No
	Homogeneity of variance-covariance matrices	23.24	.079	Yes

Results of robust tests due to violation of assumptions

Overall, no significant differences were found between the results of the parametric tests reported in the main manuscript (following the original analysis of Pronin and Ross (2006)) and the supplementary robust tests calculated due to the violation of assumptions. Please refer to the OSF for the R script used to calculate these tests. Please refer to Table S21 for a summary of the results of the parametric tests in comparison to the non-parametric tests.

Table S21

Summary of results of the parametric vs non-parametric statistical tests

Factors	Dependent variable	Parametric test			Non-parametric test		
		test-stat	*df*	*p*	test-stat	*df*	*p*
3 (temporal distance: past vs present vs future) x 2 (social distance : self vs friend) between-subjects MANOVA							
Temporal Distance*Social Distance	Total traits	*F* = 3.89	4	.004	*F* = 16.52	4	.002
Temporal Distance	Total traits	*F* = 10.65	4	< .001	*F* = 44.26	4	<.001
Social Distance	Total traits	*F* = 18.17	2	< .001	*F* = 36.34	2	<.001
Post-hoc between-subjects ANOVA for total traits (DV1) and positive traits (DV2) following significant MANOVA term							
Temporal Distance	Total traits	*F* = 0.24	1	.791	*F* = 0.90	2	.409
Social Distance	Total traits	*F* = 9.02	2	.003	*F* = 6.21	1	.013
Temporal Distance	Total positive traits	*F* = 16.88	2	< .001	*F* = 12.46	2	<.001
Social distance	Total positive traits	*F* = 36.20	1	.003	*F* = 48.86	1	0
Post-hoc independent sample t-test for self condition across temporal distance							
Temporal Distance (present vs past self)	Total traits	*t* = -0.21	287.58	.669	*t* = 0.02	NA	.975
Temporal Distance (present vs future self)	Total traits	*t* = 0.49	286.01	.623	*t* = 0.90	NA	.349
Paired sample t-test : positive vs negative traits in present-self condition							
Temporal Distance (present self)	Total positive vs total negative traits	*t* = -1.68	274.72	<.001	*t* = 8.85	87	0

Additional information about the study

1. Setting: experiment was conducted via an online questionnaire using Qualtrics. Therefore, there was no control over the physical setting in which the experiment was conducted.
2. Duration of Study Sessions: participants were allowed to take as much time to complete the online questionnaire, but were limited to complete the study for a maximum of 30 minutes. Otherwise, no time limits were set.
3. Time of Day: Since the questionnaire was completed online, there was no specific time of day as participants completed the questionnaire when convenient for them.
4. Data collection dates: The online questionnaire was opened starting 9PM on September 8, 2020, to 9PM on September 9, 2020.
5. Participant Recruitment: Participants were recruited using the Amazon MTurk platform.

Data collection procedures:

This study was conducted on Amazon Mechanical Turk with American participants. We imposed the following settings in recruiting our participants:

1. Participants were paid $0.63 as a fixed participation reward. This amount was determined by multiplying the expected completion time (in mins.) with the minimal federal wage in the U.S. (i.e., $0.125 per minute).
2. The expected completion time was set at 4 minutes in advance.
3. The most time we allowed each worker to complete the study was 30 minutes.
4. We limited all workers' HIT Approval Rate to be between 95% and 100%.
5. We limited each worker's number of HITs approved to be between 1,000 and 50,000.
6. We only used "CloudResearch Approved Participants" with Block Suspicious Geocode Locations and Block Duplicate IP Addresses enabled.
7. We blocked duplicate IP addresses and duplicate geolocation.
8. We enabled HyperBatch.
9. We restricted workers' location to be in the U.S.

References

Ben-Shachar M, Makowski D, Lüdecke D (2020). "Compute and interpret indices of effect size." *CRAN*. R package, https://github.com/easystats/effectsize.

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., ... & De Rosario, M. H. (2018). Package 'pwr'. R package version, 1(2).

Ioannidis, J. P. (2008). Why most discovered true associations are inflated. *Epidemiology*, 640-648. https://doi.org/10.1097/EDE.0b013e31818131e7

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, *1*, 389-402. https://doi.org/10.1177/2515245918787489

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, 3. https://doi.org/10.15626/MP.2018.843

Nisbett, R. E., Caputo, C., Legant, P., & Marecek, J. (1973). Behavior as seen by the actor and as seen by the observer. Journal of personality and Social Psychology, 27(2), 154. https://doi.org/10.1037/h0034779

Perugini, M., Gallucci, M., & Costantini, G. (2014). Safeguard power as a protection against imprecise power estimates. *Perspectives on Psychological Science*, 9(3), 319-332. https://doi.org/10.1177/1745691614528519

Pronin, E., & Ross, L. (2006). Temporal differences in trait self-ascription: When the self is seen as an other. *Journal of personality and social psychology*, *90*(2), 197. https://doi.org/10.1037/0022-3514.90.2.197

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.