Point-by-point Responses to the Reviewers' Comments

We would like to express our sincere gratitude for the time and effort to review our manuscript.

Reviewer 1

1.1.1: The Methods section is not detailed enough

My biggest issue which is also related to most of my other comments is that the manuscript is not transparent enough and does not include enough information on the proposed study for a Stage 1 registered report. Most importantly, the Methods section is not detailed enough to replicate the proposed study based on the given information.

All measures used in the pilot and main study should be described in more detail: The authors mention that they measure attributional confidence based on Clatterbuck (1979) and attraction based on Montoya & Horton (2004). However, this does not precisely specify the applied scales. Clatterbuck (1979) includes both a long and a short scale. I suppose the authors refer to the 7-item scale, but when reading the items in the original publication I am not sure if all items can be used without adaptation in the context of the study. The same applies for the scale in Montoya & Horton (2004), in which case the authors mention that they will use an adaptation. All items which will be actually used should be included in the Stage 1 manuscript (or in supplemental materials uploaded to the OSF repository). The authors should also make sure that the intended item response format is described in enough detail. Currently it is not clear, whether all response options will be labeled or only the endpoints of the 7-point scale. In addition to the two scales, all other questions included in both the pilot and the main study should be described with the exact wording. This includes the questions on rating the abstractness and the favorability of acts in the pilot study, all questions on demographic variables, and the concrete DQS questions. I would suggest to add a dedicated measures section to the manuscript where all this information is given and to add all item wordings and response scales to the OSF repository.

Moreover, it must be better described how the final profiles will be visually presented in the main study. I would suggest to describe the appearance of the final profiles in a dedicated section and upload all materials to the OSF. Currently, it is not clear at all how the final profiles will look like, whether they include more stimuli than the manipulated sentences described in the manuscript, whether additional information will be displayed (e.g., does the person described in the profile have a name or will be a generic picture included). The concrete instructions for the participants that will rate the profiles in the main study should also be described.

Thank you for your valuable comments. Following your suggestion, we added all scales in both the pilot and main studies and DQS questions on the Method section and the Appendix at the OSF repository (https://osf.io/69byn/files/osfstorage?view_only=dc1bb4d7647046ccae4d64ba44448921). It also contains an example image of what participants will find profiles and the actual format of the survey.

1.1.2: Important assumptions of causal mediation analysis are not discussed

Recent methods literature criticize how causal inference (e.g. mediation analysis) is performed in psychological research (e.g., Rohrer et al. 2022; Rohrer 2018), and emphasize the untestable assumptions these analyses rely on. I think it is very important that the authors

discuss the causal assumptions necessary to test the hypothesized mediation effect in the proposed main study with as little bias as possible. To my understanding, the most crucial assumption in a design such as this (where the independent variable is randomized but the mediator is not) is the absence of post-treatment confounders that have a causal effect on both the mediator and 2 the outcome. I have visualized the causal DAG of this setting in Figure 1. In the current study, the mediator and the outcome are scores where different raters rate different profiles. Thus, both rater and profile characteristics that have a causal effect on both attributional confidence and attraction would bias the estimation of the mediation effect. I think it might be possible to rule out most profile effects by experimental control of the stimuli. However, I think the absence of confounding rater characteristics is unrealistic: Without being an expert on attraction research, I could imagine a number of vaguely defined personality traits that might result in raters perceiving both less confidence in the behaviors of others and less attraction to others in general. When such rater characteristics (which cannot be ruled out and are not affected by the randomization of the independent variable) are present, the analysis might find a mediation effect even if there is no causal effect of attributional confidence on attraction. I will give a short example based on a simple simulation in the Appendix. To summarize, I think that a better description of the theoretical estimand (i.e., the investigated total and mediation effects) of the study and a discussion on how to estimate those effects most precisely would greatly strengthen the theoretical and methodological foundation of the study.

Thank you for the thoughtful comment. As you indicated, we also consider that it is crucial to assume possible confounders in our median model. Thus, we now plan to use a linear mixed model to test H1 and a multi-level mediation analysis for H2, and we will add a random effect of participants (rater) to our model to control possible confounders. In addition, we added the following sentences discussing this problem to our manuscript.

"Given the multilevel structure of the data, we will test our hypotheses using a linear mixed model (LMM) implemented in the R package "lme4" (Bates et al., 2015) and "mediation" (Tingley et al., 2014). **[...]** ***If the indirect effect becomes significant in a one-sided test (i.e., the lower bound of 90% confidence interval is above 0), we will adopt H2."***

(lines 361 - 371)

1.1.3: Aggregated data analysis might not be appropriate because of potential post-treatment confounders

The problem of post-treatment confounders in the previous comment is related to the question on whether the aggregated data analysis proposed in the manuscript is appropriate. Although the design consists of ratings made by several raters of several profiles, the authors plan to aggregate both attributional confidence and attraction by rater and perform data analysis on these aggregated scores. In contrast, it would be perhaps more appropriate to analyze the individual ratings by using multilevel models. I have put some thoughts into whether it matters which analysis to choose based on a simple simulation shown in the Appendix. I think the aggregated analysis can be used to estimate the total effect of the abstractedness manipulation on attraction. Thus, testing H1 should be possible to answer with the simple t-test proposed by the authors (although I would suspect higher power by using a multilevel model). To estimate the hypothesized mediation effect in H2, the situation is more difficult. I think the aggregated analysis should still be fine if we assume that there are no

unobserved post-treatment confounders (as demonstrated in my simulated example). In contrast, if some personality traits of raters effect both the mediator and the outcome (which I think is highly plausible), the aggregated mediation analysis should always be biased. However, I think it might be possible to control for rater characteristics (at least in some simple scenarios) by using multilevel mediation analysis. As demonstrated in my simulated data example, it seems theoretically possible to estimate the correct mediation effect despite a post-treatment confounder. I think this might be a good argument to use multilevel mediation to investigate H2 (and in that case also using multilevel models for H1 would probably be more elegant).

We appreciate your helpful advice. We now understand that analysis using aggregated data can make it impossible to control the effects of unmeasured confounders. Thus, as we mentioned above, we plan to use a multi-level model for testing H1 and H2, and control them with a random effect by participants.

1.1.4: The number of different profiles might be too small

The possibility to control for post-treatment confounding with multilevel mediation seems to rely on a sufficient number of rated profiles per person. As can be shown with my simulated example in the Appendix, when confounding is present the mediation effect can only be estimated without bias if each rater rates a large enough number of profiles. I think, the proposed number of only three different profiles will not be enough to profit from the possibility to control for confounding with multilevel models. For this reason, I suggest that the authors think about increasing the number of profiles in their main study.

Thank you for the valuable comment. As you pointed out, it is plausible that there are some variables possibly confounding the results. Thus, we decided to increase the number of profiles from three to ten. In addition, we confirmed that we can successfully detect the target effects with the ten profiles in our simulation.

1.1.5: Generalizability and ecological validity is questionable

The optimal number of profiles is also related to the question on the expected generalizability of the proposed study. Recent work has criticized that many psychological studies are not designed and analyzed in a way that justifies the intended generalizable claims (Yarkoni 2022). The current study hypothesizes that the abstractedness of profiles on social media platforms has a causal effect on how attracted people are towards the person described in the profile. However, the proposed study design relies on only three examples of such profile. Yarkoni (2022) calls this “the stimulus-as-fixed-effect fallacy”. I think it is highly plausible that even if abstractedness has on average a causal effect on attraction, this effect might vary between profiles (for some profiles the effect might even be positive). If only a small number of profiles is included in the study, there is a high chance that the average effect across the small sample lies far away from the true average effect in the population of all relevant profiles. For this reason, I think including a larger sample of profiles would increase the generalizability of the proposed study in addition to the additional advantages of controlling for post-treatment confounding mentioned in the previous section. Note that multilevel models could be used to explicitly model profiles with varying effects, when the number of profiles is not too low.

Another important issue I have with the generalizability of the proposed study in its current form is the ecological validity of the profiles presented in the main study. In their introduction, the authors state that they want to investigate the hypothesized effect of abstractedness on attraction in the context of online social media platforms. However, they do not describe the intended setting any further, which greatly limits the potential generalizability of their findings. It is not clear to me whether the virtual profiles intended to be included in the main study are representative for actual profiles on social media platforms. Because the authors currently only give a small number of examples of act descriptions and they do not describe how the complete profile will be presented, the reader cannot asses whether the study findings might be transferable to actual data from social media. Based on the current examples included in the manuscript (e.g., "I belonged to a basketball club in my school days."; "He helped an elderly person."), I am skeptical that such descriptions are representative for real online profiles on social media platforms. I think to make the study stronger, the authors should pick a concrete real live setting (e.g., a social media platform like Instagram or a dating platform like Tinder) and aim to construct their profiles as representative for the chosen setting as possible (without disregarding important aspects of experimental control). I want to emphasize again that the concrete materials and a visual example of how the final profiles will be presented should be included in the OSF repository to strengthen the Stage 1 manuscript and increase both the generalizability and the reproducibility of the research.

Another relevant aspect of generalizability might be cultural differences in the psychological variables of interest (Deffner, Rohrer, and McElreath 2022). Because my knowledge on both attraction research and Japanese culture is limited, I was wondering whether Japanese people differ from people in North America or Europe with respect to how they describe themselves in social media platforms, how favorable they assess certain behaviors, or how they perceive or describe attraction towards people they meet online. A discussion on documented cultural differences on these topics might enable readers to better assess how well the findings of this study can be generalized to populations from different countries.

Thank you for your thoughtful comments. First, for the generalizability issue on our stimuli, we have addressed it by increasing the number of stimuli from three to ten. We believe that we can now approach the true effect of the abstractness more accurately than the previous plan.

Second, regarding the issue of ecological validity, the previous manuscript could be read as if the main objective was to examine the effect of the online social media profile in context. However, our most important goal in this study is to examine the effect of abstractness on attraction in a more general scene, not limited to social media profiles. Thus, in our new manuscript, we put less emphasis on a specific context and examine the effect of the abstraction with fewer confounding variables. In addition, we will add actual profiles used in the experiment in the OSF.

Lastly, regarding cultural differences, we mention this issue in the introduction section as follows.

"*Furthermore, our study addresses a cultural gap in existing knowledge concerning uncertainty.* […] *It would provide a novel insight into the role of uncertainty in a novel form of communication—browsing others' information unilaterally.*"

(lines 197 - 202)

1.1.6: Inadequate randomization to experimental conditions

The authors propose to ask the participants in the main study about their month of birth and randomize the experimental condition based on whether the month of birth is an even or uneven number. This randomization is not appropriate: First, birth months are not completely random because pregnancies can be planned (at least to a certain degree). Second, relying on such imperfect randomization is totally unnecessary because all software packages for online surveys have implemented functionality for true randomization.

We acknowledge that our randomization method was indeed inappropriate. We have decided to change an experimental platform from Google Forms to a webpage programmed by jsPsych (de Leeuw, 2015). We will use "jsPsych.randomization.sampleBernoulli" function for randomization of the experimental condition.

"Participants will be assigned randomly to either the concrete condition or the abstract condition. For randomization, we will use the "jsPsych.randomization.sampleBernoulli" function within jsPsych."

(lines 315 - 317)

1.1.7: Inadequate use of two-sided hypothesis tests

The authors propose two-sided hypothesis tests throughout their manuscript although they have directed theoretical hypotheses (e.g., lower abstractedness is supposed to lead to higher attraction). I think the authors should always perform one-sided tests in a preregistered study with directed hypotheses.

Thank you for your suggestion. We have changed a two-sided test to a one-sided test in our analysis plan and re-calculated the sample size based on this change.

1.2.1: Sample size justification and description of mediation analysis is not detailed enough

For H1, the authors provide a sample size justification based on power analysis. Although I could reproduce their result in R (see below), the R code should be included in the OSF repository to show how sample size was calculated.

For H2, the authors cite another mediation study to justify their sample size:

"For H2, as a previous study showed that a sample size of 462 was needed to detect a small mediation effect with 1 - β = .80 (Fritz & MacKinnon, 2007), we decided the number of participants as 500."

Although I might agree that a sample size of 500 raters seems enough (at least if the number of profiles is increased) I think this justification is too weak and should be revised (see for example, Lakens 2022).

The authors should describe exactly how they plan to run their analyses in more detailed. Especially for the mediation analyse, the mediation package has several options which were not specified in them manuscript. I suggest to include the R code that will be used to investigate H1 and H2 in the OSF repository.

Thank you for your comment. We re-performed the sample size calculation using the Monte Carlo simulation. The analysis showed that 880 participants are needed to detect the target effect with a power(1-β) of .80. We registered the R code used in the simulation (https://osf.io/69byn/files/osfstorage?view_only=dc1bb4d7647046ccae4d64ba44448921).

1.2.2: Discussion of manipulation tests of the pilot study

The authors propose to test in the pilot study whether i) the rated abstractedness differs, and ii) the rated favorability of acts does not differ between sentences describing persons. However, the authors do not really discuss why these checks are performed and how they are connected to the theoretical assumptions of their causal model. Aspect i) is straightforward and seems like a reasonable manipulation check to ensure that the abstractedness manipulation works and thus abstractedness has the potential to influence attributional confidence (i.e., increase the plausibility that path $a > 0$). Aspect ii) seems less clear to me. One might argue that favorability of acts is another mediator besides attributional confidence. By removing profiles that differ with respect to this potential other mediator, the study will be designed in a way to increase the ratio of the total effect which can be mediated by attributional confidence. I am also not sure how favorability of acts is related to attraction in the theoretical model. Both constructs seem somewhat similar and the authors should discuss the theoretical differences. The potential problem I see is that the favorability ratings and attraction ratings might measure more or less similar constructs. If this were true, including only sentences with comparable favorability would lead to a reduction of the total effect to (almost) 0, which would ruin the intention of the main study. I think a deeper discussion of these aspects might be useful to strengthen the manuscript by clarifying the rationale of the manipulation tests of the pilot study.

Thank you for your suggestion. First, as you indicated, the purpose of measuring the level of abstraction is for manipulation checks. We will confirm that the abstractness of stimuli differs significantly in the expected direction.

On the other hand, the purpose of measuring favorability is to control the effect of the information's favorability, which can be confounded with the effect of abstractness, and extract only the effect of uncertainty reduction which we are interested in in this study. In this experiment, we aim to assess the effect of abstractness on a target person's attraction. Thus, we need to control variables other than abstractness that can affect attraction.

Given information about a target person, the attraction of the person is not determined only by the sum of the favorability of each piece of information. Once people receive information about a target person, they try to integrate them to construct a representation of the target and form an impression including attraction, based on that representation. We predict that the effect of the abstractness of expression (i.e., uncertainty) works in the latter part of this process. In other words, even if one receives information with the same favorability, more concrete information will encourage them to form a more favorable impression of the target person, because it reduces uncertainty of the target. Thus, even if we control the favorability of the action *per se*, we expect that profiles expressed less uncertainly, that is more concrete, are connected to higher attraction by reducing the uncertainty of the target person.

We added the above discussion to the preliminary survey section, as follows.

"*To mitigate potential confounding variables beyond abstractness, we intentionally choose sentence pairs with similar favorability.* [...] *Thus, even when controlling for the inherent favorability of the action itself, we expect that profiles expressed with less uncertainty (i.e., more concrete), foster heightened attraction by minimizing uncertainty of the target person.*"

(lines 218 - 229)

1.2.3: Include manipulation check in main study

Although the authors plan to test their stimulus material in their pilot study, it might be a good idea to also include a short manipulation check at the end of the main study, to ensure that the abstractedness manipulation worked in the final setting. As I understand it, the complete profiles and how the sentences are displayed will differ between the pilot and the main study.

Thank you for your suggestion. We added manipulation check questions asking participants' perceptions of the abstractness of the profiles.

1.2.4: Literature section is a bit short

The introduction section discussing the literature around the hypothesized effect seems a bit short. Although I am no expert on attraction research, I would suspect that there are also theories that postulate that more vague information about a person might also lead to increased attraction (e.g., if the described person is perceived as "mysterious"). A more detailed introduction will strengthen the manuscript if it discusses the exact setting in which the uncertainty reduction theory should apply and make sure that the proposed study is as representative for this setting as possible.

More explanations and literature were added to support our proposal. As you pointed out, vague information might also increase the target person's attraction. For instance, Norton and colleagues (2007) suggested more vague information about the target person increases the person's attraction. However, as we mentioned in the manuscript, there are some criticisms about their experiment. In addition, the advocates of the URT explained that the uncertainty reduction theory is adaptable mainly at the early stage of interpersonal relationships (Berger & Calabrese, 1975). Because the setting of our experiment deals with attraction toward others with no prior interaction (i.e., the early stage of relationships), we think that referring to the uncertain reduction theory is reasonable. We added some discussion on this topic as follows.

"*It is plausible that excessive certainty can dampen interest and fail to engender attraction in more mature relationships.* [...] *However, in the early stages of a relationship—when uncertainty prevails—it is reasonable to posit that uncertainty reduction will contribute to heightened attraction before people get bored.*"

(lines 124 - 131)

1.2.5: Some sentences are hard to understand

The following sentences were hard for me to understand, or contain grammatical mistakes:

- "The authors concluded that the increase in information decreased attraction since it clarified unfamiliarity, although it dissolved dissimilarities."
- "Contrarily, Reis et al., (2011) pointed out that the selection of items by Norton et al. in their study, deviated from the existing situation and conducted an experiment depicting a real interaction situation."
- "In their experiment, participants were paired and asked questions which written on cards to each other, as in a daily conversation."
- "In the main experiment, we will examine the effect of the abstractness of expression in a sentence that describes the target person's act on their attraction."
- "In addition, the act referred to in an abstract sentence semantically connotes what the opposing concrete sentence indicates."
- "We will set a significant level (α) as .05 in the analyses."
- "As d = 0.2 is considered a small effect size in behavioral sciences (Cohen, 1988), we expect that the TOST with $\Delta = 0.1$, could reduce the ff error sufficiently."

We corrected those sentences as follows.

Before	After
"The authors concluded that increase in information decreased attraction since it clarified unfamiliarity, although it dissolved dissimilarities."	deleted
"Contrarily, Reis et al., (2011) pointed out that the selection of items by Norton et al. in their study, deviated from the existing situation and conducted an experiment depicting a real interaction situation."	However, Reis et al. (2011) highlighted that the trait information paradigm overlooked the intricate dynamics of interpersonal interaction and information exchange in real-world contexts.
"In their experiment, participants were paired and asked questions which written on cards to each other, as in a daily conversation."	deleted
"In the main experiment, we will examine the effect of the abstractness of expression in a sentence that describes the target person's act on their attraction."	deleted
"In addition, the act referred to in an abstract sentence semantically connotes what the opposing concrete sentence indicates."	deleted
"We will set a significant level (α)	"We will use a significant level (α)

as .05 in the analyses.”	of .05 in the analyses.”
“As d = 0.2 is considered a small effect size in behavioral sciences (Cohen, 1988), we expect that the TOST with Δ = 0.1, could reduce the ff error sufficiently.”	deleted

Reviewer 2

2.1.1:
Flesh out the theory just a bit more. As it stands it seems implausible and that may well be because I don't understand it. Anyway, making it clearer would help. I am not an expert in this area and maybe this is plain to people in the field; but it would be good to make things clear for other interested readers as well.

Some reactions to specific claims in the paper:

"URT assumes that an individual's attraction decreases when their uncertainty increases and vice versa."Try telling that to my first wife. Apparently I was far too predictable. "Surprise me some time!" "Yes honey, how would you like to be surprised?" "Groan."

The URT theory contradicts another theory, namely that arousal of any form can potentially be turned to personal attraction. So some uncertainty could help boost arousal, and hence attraction. Morin in his 2012 book "the erotic mind" describes uncertainty as an aphrodisiac. Cindy Meston at Texas has a sympathetic arousal transfer theory of sexual arousal - e.g. doing moderate exercise helps boost subsequent sexual arousal. I am not saying to cite these particular people; there must be many papers the authors know about that are closer to their paradigm. The authors should cite some of this literature which would make opposite predictions.

Thank you for your thoughtful comment. First, we clarified what role URT plays in our idea. Specifically, we explicitly distinguished two types of uncertainty: *predictive* and *explanatory* uncertainty, and added definitions and explanations of them. Second, we included a more thorough discussion about the relationship between information amount and attraction. We added some paragraphs describing literature that can predict the opposite effect of uncertainty (i.e., uncertainty increases attraction) and explained why we can expect that uncertainty will decrease attraction in this experiment, as follows.

"*On the contrary, some studies reported that the relationship between the amount of information and attraction is not necessarily positive.* […] *However, in the early stages of a relationship—when uncertainty prevails—it is reasonable to posit that uncertainty reduction will contribute to heightened attraction before people get bored.*"

(lines 95 - 131)

2.1.2:
On p 5 there is the claim that "more information causes one to perceive more attraction to a target person at the beginning of building a relationship."
Is the theory restricted to the beginnings of relationships? If so, this should be stated upfront.

Thank you for your advice. We specified that the URT explains the beginning of human relationships as follows.

"*The uncertainty reduction theory (URT: Berger & Calabrese, 1975) explains the correlation between the amount of information and attraction. URT explains interactions in the early stage of relationships using the concept of uncertainty.*"

(lines 75 - 77)

2.1.3:

p 4 "people perceive higher uncertainty and lower attraction for others when they have different opinions and attitude"

I can certainly think of some people I am not attracted to because their disagreeable opinions are predictable.

p 4 "people perceive higher uncertainty and lower attraction for others when they have different opinions and attitude"

I didn't follow this sentence. I might know someone who is a keen Chinese Communist Party member/Trump supporter/Brexiteer/Scientologist, and find their beliefs highly predictable yet different from mine.

Thank you for your comment. We removed the relevant parts that you commented on by re-organizing the introduction section. According to URT (Berger & Calabrese, 1975), when the attitude of the target person is not similar to the rater, raters seek to explore possible attribution of the target's behavior. Thus, *explanatory* uncertainty, which is defined by the number of possible factors of the attitude or behavior of someone at interactions, will rise when the target's attitude is inconsistent with the rater. This claim is by the investigation that people are likely to attribute more causes to the person's attitude when the attitude is dissimilar than when it is similar.

2.1.4:

"deviated from the existing situation"

I didn't follow. What the authors do is have high or low uncertainty regarding positive statements. If the theory is that people are liked more when their estimate of their goodness has more precision, then this makes sense, but is a somewhat different theory. It could be seen as a version of the URT in a particular context. Some clarity on the theoretical relations here would be good. if the precision of goodness theory is found false then so is URT.

Thank you for your comment. We aim to manipulate the uncertainty of the target person by changing the abstractness of the behavior of that person. This procedure is based on the idea that abstractly expressed behavior is less informative than concretely expressed behavior. Estimation of goodness is not the main topic for this study since we believe people are less likely to estimate the goodness of the target persons in the procedure of this experiment.

2.2.1:

p 8 "If their abstractness would not differ significantly, that pair would be excluded from the candidates"

It may be better to use a threshold amount of difference e.g. significantly more than a 2 point difference.

Thank you for your advice. We decided to change the relevant procedure as follows: We will exclude pairs whose abstractness do not differ significantly or the effect size of its difference is below Cohen's *d* = 0.2. This is to ensure that our manipulation works efficiently enough.

2.2.2:
Justify N for the norming study to be sufficient to get a 90% CI within a null region of +- 0.1 Likert units. For example, find the N that will put the 90% CI inside or outside the null region a certain pecentage of times (say 80%) if H0 is true or if a plausible H1 value is true (see https://doi.org/10.1525/collabra.28202).

Thank you for your comment. We calculated the sample size for the Two One-Sided Test (TOST) using the PowerTOST package in R (Labes et al., 2024). The analysis showed that 215 participants were enough to achieve 80% power under $\Delta = 0.3$, SD = 1.0, $\alpha = .05$. Thus, we decided to collect 250 participants in the preliminary study.

2.2.3:
Design Table;

Why isn't the uncertainty reduction theory being tested? I wasn't sure why the authors thought it escaped contrary evidence.

Thank you for your comment. Our hypothetical model is that abstractness of expression influences the uncertainty of the target person and this leads to improvement of their attraction. The uncertainty reduction theory explains the latter part of this model. If the result showed that uncertainty did not correlate with attraction, it indeed seems to contradict the theory. Nevertheless, as we did not plan this study to test the uncertainty reduction theory, we did not directly include this relationship in the hypothesis. Thus, even if we obtained such contrary evidence, this will be an exploratory result, and we cannot reject the theory only from that. In addition, the situation dealt with in this study, which is participants unilaterally rate the target person without interactions, differs from the situation that URT originally assumed. Therefore, even if we obtain a contrary result, it will not necessarily mean the URT is invalid.

2.2.4:
H1: Why is the effect size chosen as d = 0.3 for power?

I don't know how subjects rated attractiveness, but let us say it is a 7 point Likert scale. How many Likert units is meaningful? PCI RR guidelines say "power analysis should be based on the lowest available or meaningful estimate of the effect size". One could look at previous studies that investigated UCT, and take the smallest difference these studies find (for a more rigorous approach see https://doi.org/10.1525/collabra.28202).

Thank you for your comment. Regarding the effect of uncertainty on attractiveness, we adopt a medium effect size based on our existing data. On the other hand, because of the lack of previous investigation, we decided to assume a medium effect size for the effect of abstractness on uncertainty.

2.2.5:
H2: presumably power = .08 is meant to be .80.

For mediation, the precise results that will lead to certain conclusions should be explicit. Would any degree of partial mediation do? How much would be meaningful? Best

would be to think in terms of raw units, and an indirect effect of the same minimal interesting effect size just previously identified for the total effect also used for the indirect effect, for simplicity. Work out power based on that.

Thank you for your comment. We corrected our typo and re-performed the power analysis for the mediation model using one simulation. Regarding the effect of uncertainty on attractiveness, we adopt a medium effect size based on our existing data. For the effect of abstractness on uncertainty, because of the lack of previous investigation, we decided to assume a medium effect size.

2.3.1:
"We will assign participants with even-numbered birth months to the abstract condition and those with odd-number to the concrete condition"

I suppose there could be personality differences depending on birth month? A GScholar search revealed titles claiming evidence for this even in specifically Japanese (e.g. https://www.cambridge.org/core/journals/european-psychiatry/article/abs/effect-of-month-of-birth-on-personality-traits-of-healthy-japanese/9CAD7C5E8898C29636FC64C7DD86EE4F)

So subjects should be assigned randomly.

We acknowledge that our randomization method was indeed inappropriate. We have decided to change an experimental platform from Google Forms to a webpage programmed by jsPsych (de Leeuw, 2015). We will use "jsPsych.randomization.sampleBernoulli" function for randomization of the experimental condition.

"Participants will be assigned randomly to either the concrete condition or the abstract condition. For randomization, we will use the "jsPsych.randomization.sampleBernoulli" function within jsPsych."

(lines 315 - 317)

2.3.2:
p 10 "Afterward, participants will complete questionnaires regarding their attributional confidence"
Describe the questionnaires. How many questions? What is the rating scale?

Thank you for your valuable comments. Following your suggestion, we added all scales in both the pilot and main studies and DQS questions on the Method section and the Appendix at the OSF repository (https://osf.io/69byn/files/osfstorage?view_only=dc1bb4d7647046ccae4d64ba44448921). It also contains an example image of what participants will find profiles and the actual format of the survey.

References

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. Human Communication Research, 1(2), 99–112. https://doi.org/10.1111/j.1468-2958.1975.tb00258.x

Norton, M. I., Frost, J. H., & Ariely, D. (2007). Less is more: the lure of ambiguity, or why familiarity breeds contempt. Journal of personality and social psychology, 92(1), 97–105. https://doi.org/10.1037/0022-3514.92.1.97

Labes, D., Schütz, H., & Lang, B. (2024). PowerTOST: Power and Sample Size for (Bio)Equivalence Studies. R package version 1.5-6. https://CRAN.R-project.org/package=PowerTOST