**Peer Review History and Communication**

**Ms Title:** Prejudice at First Sight? First Impressions of Refugees are More Strongly Influenced by Targets’ and Perceivers’ Interindividual Characteristics Than by Sheer Group Affiliation

**Author Names:** Joscha Stecker, Paul C. Bürkner, Jens H. Hellmann, Steffen Nestler, & Mitja D. Back

**Submitted:** July 23, 2020

**Editor First Decision—Revise & Resubmit**

August 7, 2020

Dear Joscha Stecker,

Thank you for submitting your work to Collabra: Psychology under the streamline review mechanism. I read the paper and then the letter of response to the prior reviews. The paper is detailed, and the reporting is thorough. I think there is a possibility that this work would be a good fit for Collabra. However, I think the current framing of the paper is misaligned with the empirical work and this makes it challenging to fully determine the suitability of the paper for Collabra. A significant issue is whether the claims in the paper are appropriately calibrated to the design and statistical evidence in the work. Thus, my decision is to offer you a chance to revise the work in light of my reactions. If you do elect to revise the work, I will send the paper out for review to experts on person perception especially perceptions that relate to stigmatized groups. To my mind, there are open questions about the internal validity and implications of the work (see below). I would like some additional expert review about these issues. I just want to get the paper in a place where I believe peer review would be useful for my decision making purposes.

Below I outline what I think would make the paper suitable for review.

1. My reading of the comments of the previous editor and reviewers was that they had a major concern about how this work advances scholarship on prejudice in the ways the topic was typically studied in the social psychological literature. I am not an expert on this topic, but I think this criticism still applies to this revision. In other words, I did not believe the revised paper adequately addressed their concerns. I will try to articulate my perspective and then offer a potential pathway for revision.

At the end of the day, I did not find the null result for overall devaluation to have much relevance as to whether and how strongly people hold prejudicial attitudes toward refugees or minoritized groups. The particular task faced by participants in these studies was not especially suited for addressing such a big picture question. I have no problem with null results, and I think null results can be informative. However, not all null results are informative with respect to particular claims.

To address this issue, I think the paper should be framed almost exclusively around person perception. I think this will situate the paper in the appropriate literature and this pruning could also help reduce the length of the paper. It would be fine to speculate briefly on connections with the social-psychological literature on prejudice in the Discussion section, but I think this material should be expressed more tentatively.

I also think some consideration of whether social desirability pressures in a lab study are relevant to the interpretation of these results would be valuable. This process could also chip away at a substantive interpretation of the null findings.

1. In line with #1, I also think more attention should be paid to the lack of labels accompanying the photographs. I worried that this was a greater threat to internal validity than was acknowledged in the revision and response to prior reviews. I think it is largely an untested assumption that participants recognized group membership during the task. (If I missed some counterevidence in the paper, I am sorry). This issue could be magnified in Study 2 if the religious labels somehow overshadowed the minoritized versus majority group distinction implied by the photographs. Footnote #4 might be less informative than it would appear as the judgement task itself could have prompted a particular social categorization whereas it is not so clear to me what social cognitive processes were actually activated in the actual task facing participants.

In retrospect, it might have been useful to sprinkle some manipulation check items within the task itself. However, I also respect that this could have introduced demand characteristics. I guess one option would have been to ask participants to also rate targets on the likelihood that they experienced discrimination from others (or something more neutrally framed). Alternative designs are possible where even brief person descriptions paired with photographs are contrasted with photo only and person description only conditions. At any rate, I think a revised version should grapple with this issue and perhaps even have a section in the Discussion that describes how the results should be interpreted if this reservation about the lack of labels has merit.

1. It would be useful to address whether and how participants were debriefed given that some deception was involved (e.g., the faces were literally pictures of Americans rather than Germans). Likewise, the faces were from particular databases and it would be important to verify that the use of these faces was consistent with the original understanding of use from the people providing the photographs.
2. Please include a formal constraints on generalizability section in the Discussion of any revision (Simons et al. 2007).
3. Is there a way to implement equivalence tests into the analysis/interpretation to evaluate if the parameter estimates are inconsistent with even tiny effect sizes? Alternatively, is there a way to consider the reported 95% CIs in ways that would work with the spirit of equivalence tests? I also thought that effect size discussions could be a bit more prominent in the paper.

6 Please discuss “satisfactory interrater agreements” in just a little more detail such as 1 or 2 sentences. This could include reference to typical values in the field. Did agreement differ for the Iranian versus American faces?

1. In general, I think the paper is quite long and it has 20 pages or so of references. On the one hand, I appreciate the thoroughness of the scholarship. But I do think the sheer size of the paper might strike some readers as off-putting. I try to avoid micro-managing issues of style when I handle papers, but I thought this comment could prove helpful if you revise the work. If you follow my point #1, some length will be removed from the paper. I think this will further highlight tthe empirical contribution of the work to the person perception literature.

In summary, I think the task in a revision is to reconsider the framing of the paper and to engage even more with the previous criticisms of the work. I think revising along these lines will make the paper suitable for further review at Collabra. I hope this feedback is useful even if it might not be so welcome. You may decide that this work will face an easier time to publication at another outlet. I wanted to be upfront about the issues so you can make the most appropriate decisions for your team. If you have any questions or difficulties during this process, please contact the editorial office at [editorialoffice@collabra.org](mailto:editorialoffice@collabra.org).

Thank you for considering Collabra and again I hope this feedback proves useful regardless of the fate of this paper at this outlet.

Sincerely,

Brent Donnellan

**Author Response**

Dec 17, 2020

 Dear Prof. Donnellan,

Please find attached a revised manuscript entitled “First Impressions of Refugees are More Strongly Influenced by Target Cues and Perceiver Attitudes Than by Sheer Group Affiliation” which we would like to resubmit for publication in Collabra: Psychology.

With the help of your revision remarks, we rewrote several parts of the previous manuscript and now focused on person perception with insights on first impressions of refugees. In the paper, we assess both perceiver and target characteristics to better understand determinants of first impression outcomes. Hereby, the design with individual target photos allowed us to differentiate between the effects of individual target cues (e.g., smiling), target group affiliations (e.g., Muslim) and perceiver characteristics (e.g., SDO) on perception outcomes. Thereby, the presented study helps to better understand the perception of refugees as one of the most heatedly discussed minorities in Western societies. Across two studies, we provide evidence for (a) no overall devaluation of refugees or Muslims, (b) strong effects of target attractiveness and smiling on evaluations across individuals of different group affiliations, (c) strong effects of perceiver attitudes towards refugees and Muslims, and (d) no interactive effects of perceiver attitudes and target cues on evaluations. These findings highlight the utility of an individual difference approach to understand perception outcomes of refugees and show first evidence for differences between perceiver and target determinants. We believe this manuscript will be of great interest to many readers of Collabra:Psychology.

This paper has been first submitted to JPSP and has been rejected after reviews. Afterwards, it was submitted in a streamlined process to Collabra:Psychology and was sent back for revision. Below we summarize, how we have incorporated your editorial feedback to make the paper applicable for revision.

Please note, that we widely revised the work and even change the manuscript title to better match our focus on person perception and different determinants of first impression outcomes of refugees. Thereby, we aim to contribute to a better understanding of the different characteristics on perceiver’s and target’s sides and their potential interactions for impression formations. We hope to offer new perspectives on the underrepresented sample of refugees within a person perception framework and show the utility of extending research on refugees to different angles (such as facial perception studies). The presented results can be an important step to further understand the impact of individual target characteristics (as opposed to group affiliations) on perception outcomes and offer first insights to perceiver’s personality and attitudes effects within person perception settings.

None of the findings have been published or are under consideration elsewhere, including the Internet.

Selected variables of the online survey for the assessment of individual characteristics in both contained studies have been analyzed in an unrelated submitted paper on threat towards immigrants that is currently under review (Bollwerk, M., Schlipphak, B., Stecker, J., Hellmann, J. H., Echterhoff, G., & Back, M. D. (2019). Feeling Threatened by Immigrants: The Role of Ideology and Subjective Societal Status. Manuscript submitted for publication). This paper does not include any of the judgment data based on target photographs that are key dependent variables in the present submission. It focuses on subjective social status as the main predictor. The only overlap refers to the inclusion of SDO and RWA as additional predictors that are also included in the present investigation.

Data were collected in a manner consistent with ethical standards for the treatment of human subjects, and APA ethical standards were adhered to during the research project.

We have read and adhered to Collabras Transparency and Openness policy. Supplemental materials including codebooks, data, R-script as well as supplementary analyses and results can be found here: http://go.wwu.de/-ygrf (shortened link to an OSF project). We refer to these materials with an anonymized view-only link in the manuscript. The paper best fits into Collabras “Personality Psychology” section as an “Original research report”. We are opting for open peer review. Please address correspondence to: Joscha Stecker, Department of Psychology, University of Münster, Fliednerstr. 21, D-48149 Münster, Germany; Phone: +49 251 8334122; Fax: +49 251 8331331; Email: joscha.stecker@wwu.de. Contact information of the other authors are as follows: Joscha Stecker (joscha.stecker@wwu.de), Jens Hellmann (jens.hellmann@uni-muenster.de), Steffen Nestler (steffen.nestler@wwu.de) all University of Münster (Germany); Paul Bürkner (paul.buerkner@gmail.com), Aalto University (Finland). We hope that you will evaluate our manuscript as suitable for publication in *Collabra:Psychology* and thank you very much for considering it. Yours sincerely Joscha Stecker (on behalf of all authors)

**\*\*Editorial remarks to our previous submission to Collabra along with our responses\*\***

**Editorial remarks:**

**The paper is detailed, and the reporting is thorough. I think there is a possibility that this work would be a good fit for Collabra. However, I think the current framing of the paper is misaligned with the empirical work and this makes it challenging to fully determine the suitability of the paper for Collabra. A significant issue is whether the claims in the paper are appropriately calibrated to the design and statistical evidence in the work. Thus, my decision is to offer you a chance to revise the work in light of my reactions. If you do elect to revise the work, I will send the paper out for review to experts on person perception especially perceptions that relate to stigmatized groups. To my mind, there are open questions about the internal validity and implications of the work (see below). I would like some additional expert review about these issues. I just want to get the paper in a place where I believe peer review would be useful for my decision making purposes.Below I outline what I think would make the paper suitable for review.**

**My reading of the comments of the previous editor and reviewers was that they had a major concern about how this work advances scholarship on prejudice in the ways the topic was typically studied in the social psychological literature. I am not an expert on this topic, but I think this criticism still applies to this revision. In other words, I did not believe the revised paper adequately addressed their concerns. I will try to articulate my perspective and then offer a potential pathway for revision.**

Thank you for the positive evaluation of our work and for helping us realizing that the framing of our paper was indeed a bit misaligned with what we are offering. We found this feedback to be extremely helpful and have now much more clearly focused on first impression as a highly relevant phenomenon in the context of refugee evaluation. Our findings provide important insights for this phenomenon independent of how they relate to previous work on prejudice towards minoritized groups. We still believe that our work also has implications for research on prejudice more generally (e.g., pointing to the fact that those results can not easily be generalized to the perceptions of individuals which is, however, regularly implied in respective papers) but have now moved this material to the discussion (where it indeed belongs).

**E1: At the end of the day, I did not find the null result for overall devaluation to have much relevance as to whether and how strongly people hold prejudicial attitudes toward refugees or minoritized groups. The particular task faced by participants in these studies was not especially suited for addressing such a big picture question. I have no problem with null results, and I think null results can be informative. However, not all null results are informative with respect to particular claims. To address this issue, I think the paper should be framed almost exclusively around person perception. I think this will situate the paper in the appropriate literature and this pruning could also help reduce the length of the paper. It would be fine to speculate briefly on connections with the social-psychological literature on prejudice in the Discussion section, but I think this material should be expressed more tentatively.**

We highly appreciate the comment on the paper’s framing and undertook it a thorough process of revising it. Following your advice, we now limited the focus on person perception of refugees and its implication. Hereby, we are able to shine a light on person perception of refugees as an important, underrepresented sample and especially on the determinants of perception outcomes. To better focus on the person perception framework, we revised the title, abstract, introduction and discussions of the paper. For the connection of relevant perceiver characteristics in intergroup setting, we point out the link to prejudice literature without claiming prejudice-related insights in the present paper. Here, it now reads on page 7:

“*In the context of evaluations of refugee individuals, and thus, evaluations of members of minoritized groups, related esearch on intergroup perception and prejudices could be particularly beneficial to understand such perceiver differences in evaluation.*”

In the discussion section, we now have a part in which we discuss a potential transfer to prejudice research. For instance, on page 57 it now reads:

*“With regards to target differentiation, future research should include further measures of targets personalities. Following lens model frameworks (Hirschmüller, Egloff, Nestler, & Back, 2013; Nestler, & Back, 2013), this would allow to additionally explore how much physical and behavioral cues are indicators of stable personality characteristics (i.e., to analyze cue validities) and mediate personality-prejudice and personality-personality impression (i.e., accuracy) associations.”*

The focus on person perception also helped to significantly reduce the length of the paper and the reference list as further pointed out under E9.

**E2: I also think some consideration of whether social desirability pressures in a lab study are relevant to the interpretation of these results would be valuable. This process could also chip away at a substantive interpretation of the null findings.**

We agree that social desirability always needs to be included in the interpretations of lab study results and therefore now added a section in the discussion highlighting this. However, we argue that the presented study design offers a reduce social desirability impact. On page 54f. in the discussion, it now reads:

“*As for most evaluative studies, social desirability has to be taken into account when interpreting results. However, it can be argued that the presented studies reduced the social desirability by (a) not asking for an overall judgment (i.e., likeability) but for certain dimensions (e.g., trustworthiness, competence, modernity) and (b) focusing on photos of individuals with certain characteristics as opposed to an overall judgment of groups (e.g., refugees or Muslims). By offering a design assessing impressions towards group (members) in a more indirect way, we reduced social desirability effects present in other studies of intergroup contexts. Furthermore, in contrast to an overall social desirability account, we found an overall religiousness devaluation effect and particularly negative attitudes towards devout Muslim.*”

**E3: In line with #1, I also think more attention should be paid to the lack of labels accompanying the photographs. I worried that this was a greater threat to internal validity than was acknowledged in the revision and response to prior reviews. I think it is largely an untested assumption that participants recognized group membership during the task. (If I missed some counterevidence in the paper, I am sorry). This issue could be magnified in Study 2 if the religious labels somehow overshadowed the minoritized versus majority group distinction implied by the photographs. Footnote #4 might be less informative than it would appear as the judgement task itself could have prompted a particular social categorization whereas it is not so clear to me what social cognitive processes were actually activated in the actual task facing participants.**

We appreciate the comment on labeling of the target’s group affiliations, which was also discussed in the previous streamlined review. Therefore, we took the opportunity to further point out extended considerations on this regard. While the precautions for interpreting results without individual labels are very important, we would like to provide several indicators showing a sufficient categorization accuracy. First, both studies introduced the target groups as individuals from one of both groups:

Middle Eastern refugees or non-migrant Germans. In the introduction text, the term refugee was defined and both social categories were pointed out and highlighted for the participants. Therefore, a labeling was set even though, we decided to not label each target individual for social desirability reasons as pointed out below. Second, therefore, we argue that the social cognitions in the categorization task of the additional study, should not differ significantly from the implicit categorization task in both main studies. Therefore, we are convinced that the additional study on target’s ambiguity provides first evidence for certain accuracy in our studies and highlighted it accordingly on page 13 and the footnote 6. Third, we argue that we replicated several findings on perceiver characteristics (e.g., SDO, RWA) known to predict devaluation effects towards minorities, but not only for the individually labeled religion and religiousness outgroup individuals (i.e., Muslim, devout individuals), but also for the not individually labeled refugees individuals. Those effects are found both for Study 1 and Study 2 and offer another strong evidence that categorization processes to the category “Middle Eastern refugee” as opposed to “non-migrant Germans” have worked sufficiently. We further highlighted this in the discussion on page 56f., which now reads:

“*Furthermore, future research should examine to what extent the type of presentation affects perception outcomes by varying the labeling of target photo (e.g., labeling refugees). While a more salient indication of the targets’ refugee status should elicit more category-based processes (Fiske, Neuberg, Beattie, & Milberg, 1987), the perception of individual stimuli with ambiguous group affiliations as in the present study is more ecologically valid for first impressions in daily life, which rarely include explicit labeling. With additional analyses on subsets of targets rated as particularly prototypically, we found evidence that ambiguous categorizations did not solely drive perception outcomes. Furthermore, perceiver’s RWA, SDO and political ideology were found in both studies to moderate devaluation effects of not only Muslims and devout individuals (both labeled), but also of refugee individuals (not labeled). As those attitudes were repeatedly found to moderate devaluation effects of minority groups (e.g., Craig & Richeson, 2014; Dunwoody & McFarland, 2018; Echebarria-Echabe & Fernández-Guede, 2006), this provides further evidence that the categorization processes of Middle Eastern individuals to the minoritized group and those of Western origin to a majority group worked in our design. Nevertheless, we cannot exclude the possibility that perceivers in our studies showed a less than perfect differentiation of group affiliations.*”

**E4: In retrospect, it might have been useful to sprinkle some manipulation check items within the task itself. However, I also respect that this could have introduced demand characteristics. I guess one option would have been to ask participants to also rate targets on the likelihood that they experienced discrimination from others (or something more neutrally framed). Alternative designs are possible where even brief person descriptions paired with photographs are contrasted with photo only and person description only conditions. At any rate, I think a revised version should grapple with this issue and perhaps even have a section in the Discussion that describes how the results should be interpreted if this reservation about the lack of labels has merit.**

Connecting to the comment above, we used this editor’s comment for an updated section in our discussion. As the presented paper should be seen as a first step to the better understanding of determinants of first impressions towards refugees, we argue that the presentation mode and design characteristics can also be one of those determinants. The presented studies focus on varying different levels of group affiliation and on the use of natural variation of individual cues. Therefore, we combined self-report with other-perceptions and independently coded facial cues. However, future studies are well-advised to extend these variations to better understand determinants such as labeling, framing and other design parameter. We argue, that our paper offers a solid first approach to highlight the importance of person perception approaches in the research on refugees and other underrepresented samples. The updated section in our discussion on page 57 now reads:

*“Future research is well advised to replicate the present results with different ways of experimentally varying group affiliation (e.g., contrasting brief person descriptions paired with photographs with photo only and person description only conditions). Another option would be to include some manipulation check items within the task itself while being careful to not include demand characteristics at the same time (e.g., asking participants to also rate targets on more neutrally framed aspects such as the likelihood that they experienced discrimination from others). More generally speaking, future research might vary both the abstractness (individual level to group level) and the explicitness of a target’s group affiliation in a more fine-grained way to systematically analyze effects on general, cue-based, and attitude-based evaluations.”*

**E5: It would be useful to address whether and how participants were debriefed given that some deception was involved (e.g., the faces were literally pictures of Americans rather than Germans). Likewise, the faces were from particular databases and it would be important to verify that the use of these faces was consistent with the original understanding of use from the people providing the photographs.**

Regarding the need for debriefing of participants, we argue that there was no excessive deception involved in our studies. While the introduction manipulated the group affiliation of the subsequent targets (in terms of religion, religiosity and ethnicity), there were no further misleading attributions to any targets. Therefore, we showed photos of individuals with incorrect labeling of their group affiliation, but did not in any way influenced the impression formation to any of these groups. Accordingly, one can assume that participants were not affected in their attitudes towards any of the depicted groups. However, we used a short debriefing at the end of Study 1, which reads: “*The aim of this study was to investigate the influence of cultural background and status on photo ratings*". For Study 2, we extended the debriefing, which reads: “*The aim of this study is to examine differences in the first impressions of German and refugee persons of Christian and Muslim faith. To this end, we want to analyze possible influences that may affect interpersonal judgements. Potential influences from previous research may include religion, religiousness, ethnicity and numerous personality traits and attitudes*.”.

Regarding the use of photo databases, we used the opportunity to double check the agreement forms and user guidelines for both databases. Neither database (IFDB or FERET) prohibited the use of photos for our studies. Both databases were requested beforehand and thoroughly reviewed for possible limitations of use with regard to our studies.

**E6: Please include a formal constraints on generalizability section in the Discussion of any revision (Simons et al. 2007).**

Following this valuable suggestion, we included a constraints on generality section in our discussion. Since the individual contents of our previous "Limitations and Future Prospects" section differ only slightly from the new "Constraints on Generality" section, we have integrated both into a new section called "Constraints of Generality, Limitations and Future Prospects". The new section added constraints on generality regarding the general design, participants, stimuli, procedure and temporal specificity. Please find the new section on page 53ff..

**E7: Is there a way to implement equivalence tests into the analysis/interpretation to evaluate if the parameter estimates are inconsistent with even tiny effect sizes? Alternatively, is there a way**

**to consider the reported 95% CIs in ways that would work with the spirit of equivalence tests? I also thought that effect size discussions could be a bit more prominent in the paper.**

Using Bayesian CIs, one can create equivalence tests by defining a region of practical equivalence (ROPE) around a point null hypothesis (zero in most cases). Such a ROPE would then define an interval null hypothesis and one could decide hypotheses as follows:

- If the CI does not overlap with the ROPE, we can \*reject\* the null hypothesis

- If the CI is completely within the ROPE, we could \*accept\* the null hypothesis.

- Else, if there is overlap between ROPE and CI, we could say the evidence is \*inconclusive\* and we can neither accept nor reject any hypothesis.

This implies that anyone seeing our results can perform an equivalence test by creating their own ROPE and comparing it to the CIs we reported. Of course, ROPEs need to be motivated by subject matter knowledge and be appropriate for the scale of the parameter(s) to be tested. Since, for a lot of the investigated effects, there is not yet enough theory to define what an effect should be considered practically equivalent to zero, we are hesitant to apply these kinds of equivalence tests in our paper. However, we now go into more detail on the possibility of equivalence testing in the discussion of Study 2. On page 46f., it now reads:

“*The context and evaluation of effect sizes and their magnitude is a challenging topic for person perception studies including both cue-level and group-level characteristics. For comparison of effect sizes, there are ways to define a region of practical equivalence (ROPE) around a point null hypothesis. This interval could then be compared to the confidence interval and help deciding about the null hypothesis. However, the definition of such a ROPE is dependent on the subject and scales of testing and therefore need to be informed by theoretical considerations or previous results. Due to the limited number of comparable studies and novelty in the comparison of cue-level and group-level characteristics, we are hesitant to define ROPEs, but would like to encourage future research to engage in the delineation of such ROPEs for different contexts and based on sufficient empirical studies.”*

**E8: Please discuss “satisfactory interrater agreements” in just a little more detail such as 1 or 2 sentences. This could include reference to typical values in the field. Did agreement differ for the Iranian versus American faces?**

The section on satisfactory interrater agreement is now updated and includes the reference of similar studies coding smiling and attractiveness as facial cues with comparable agreement indices. Furthermore, we added a section on differences in the agreement between faces of both databases. As expected, it was found that the intergroup agreement was higher for "smiling" than for "attractiveness" as a more culture-bound cue. On page 16, it now reads:

“*For each cue, we averaged the ratings across all three coders and found satisfactory interrater agreements for smiling, ICC (2, k) = .84, and attractiveness, ICC (2, k) = .77 (see e.g., Hirschmüller, Egloff, Nestler, & Back, 2013; Naumann et al., 2009). The agreement on smiling did not substantially differ between Middle Eastern targets, ICC (2, k) = .86, 95% CI [.74, .93], and Western targets, ICC (2, k) = .83, 95% CI [.69, .91], while the one for attractiveness for Middle Eastern targets, ICC (2, k) = .68, 95% CI [.41, .84], was descriptively lower than the one of Western targets, ICC (2, k9 =*

*.82, 95% CI [.67, .91].*”

**E9: In general, I think the paper is quite long and it has 20 pages or so of references. On the one hand, I appreciate the thoroughness of the scholarship. But I do think the sheer size of the paper might strike some readers as off-putting. I try to avoid micro-managing issues of style when I handle papers, but I thought this comment could prove helpful if you revise the work. If you follow my point #1, some length will be removed from the paper. I think this will further highlight tthe empirical contribution of the work to the person perception literature.**

With the clearer focus on person perception as outlined in E1, we were able to significantly shorten the paper. Particularly due to our intensive multimethodological approach (e.g., cue-based processing, group affiliation variation and perceiver questionnaires) the paper still has some length. However, the overall length is now reduced from 91 pages to 84 pages, whereby the reference list is reduced by four pages (from 21 to 18). We therefore hope the paper will be accessible for more readers.

**E10: In summary, I think the task in a revision is to reconsider the framing of the paper and to engage even more with the previous criticisms of the work. I think revising along these lines will make the paper suitable for further review at Collabra. I hope this feedback is useful even if it might not be so welcome. You may decide that this work will face an easier time to publication at another outlet. I wanted to be upfront about the issues so you can make the most appropriate decisions for your team. If you have any questions or difficulties during this process, please contact the editorial office at editorialoffice@collabra.org.**

Thanks to your comments and advice, we have sharpened the focus of our paper and provide a much more clear-cut conceptual, methodological, and empirical contribution. Now the paper more clearly and exclusively focuses on first impressions as a fundamental phenomenon to understand reactions towards refugees (and members of other minoritized groups) in real-life. In doing so, the paper now has a more selective claim on its implications and generalizability, providing important first steps to stimulate future research on person perception towards (individuals of) minoritized groups.

**Editor Second Decision—Revise & Resubmit**

March 3, 2020

Dear Joscha Stecker, Thank you for submitting the revised version of your manuscript, "First Impressions of Refugees are More Strongly Influenced by Target Cues and Perceiver Attitudes Than by Sheer Group Affiliation.” I carefully read the new revision and letter of response. I found most everything to be thoughtful and I believed that I was in a position to issue a decision without sending it out for review in light of the previous decision letters. Below I offer some reactions that I hope you will consider. I will not send the final revision out for review and plan to make a speedy final decision pending the revisions and letter of response.

Remaining Issues

1. I think the framing on first impressions of faces and the Constraints on Generality Statements help address many of my issues from the previous decision. However, I want to reduce the chances that the current findings are misinterpreted or taken out of context given the political and social realities of 2021 and the importance of sensitivity to the underlying issues. Only a few things remain in this regard.

1a. Would you be open to modifying the Title to “First Impressions of Faces of Refugees are….”?

1b. You might sprinkle even a few references to first impressions of faces or photographs throughout the paper. I think this is a bigger issue in the Discussion than the Introduction at this point. For example, on page 49 in the first paragraph, I think focusing on photographs is important. So it might read something like “we investigated perceptions and evaluations of refugee faces from photographs, that is…” Again, I do not want to be heavy handed but I am concerned about how people might over interpret the findings from this work. Moreover, being explicit and spelling out what cannot (or even should not in your opinion) be interpreted from the work might also be helpful. I think there is a space for one or two more sentence perhaps even in the Abstract as well.

1. The paper is long. That might not be avoidable but I would just ask you to consider any last cuts to make this streamlined. I caught a few typos as well so check for those as well. For example, I think “they” is missing from Footnote #2.
2. I still missed having some grounding for interpreting effect sizes here (I do identify as a personality researcher, after all!). I understand the complexities with the modeling approach and lack of obvious standardization. Thus, this might be as good as it gets but just consider this issue in the final revision.

That is it. I am sorry for the delay in issuing this decision. It was an oversight on my part in how I read the dashboard with the new (and much improved) submission site. I will do my best to issue a quick final decision.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This is the last opportunity for major editing, therefore please fully check your file prior to re-submission.

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible. We will of course be understanding given the pandemic and the extra stresses everyone is facing. Thanks for doing such a thorough job with the revision and being responsive to feedback. I hope you have found the process to be collaborative. If you have any questions or difficulties during this process, please contact the editorial office at [editorialoffice@collabra.org](mailto:editorialoffice@collabra.org).

Sincerely,

Brent Donnellan

**Author Response**

Mar 29, 2021

 Dear Prof. Donnellan,

We thank you very much for your guidance, and the time and effort invested in the manuscript. Naturally, we were also happy about your positive evaluation of our manuscript’ revised version. Your final feedback offered us an opportunity to further sharpen the scope, generalizability, and limitations of our findings. Also, we are thankful for the indeed very collaborative process and for the prospect of a timely final decision without further reviews. In the following we summarize how we have addressed all of the remaining points that you mentioned.

**E1: I think the framing on first impressions of faces and the Constraints on Generality Statements help address many of my issues from the previous decision. However, I want to reduce the chances that the current findings are misinterpreted or taken out of context given the political and social realities of 2021 and the importance of sensitivity to the underlying issues. Only a few things remain in this regard.**

We very much agree with the need to ensure a thoughtful process to minimize the threat of misinterpretation or instrumentalization of presented results for a political ideology, especially in the current societal climate. Therefore, we found the comment very helpful to go through the manuscript again and clarify the scope of research and to anticipate false interpretations of the presented results.

**E2: Would you be open to modifying the Title to “First Impressions of Faces of Refugees are….”?**

From our perspective, the proposed modified title helps to not only specify the actual research methodology, but also to prevent misinterpretations on the generalizability. Therefore, we gladly adapted the title accordingly.

**E3: You might sprinkle even a few references to first impressions of faces or photographs throughout the paper. I think this is a bigger issue in the Discussion than the Introduction at this point. For example, on page 49 in the first paragraph, I think focusing on photographs is important. So it might read something like “we investigated perceptions and evaluations of refugee faces from photographs, that is…” Again, I do not want to be heavy handed but I am concerned about how people might over interpret the findings from this work. Moreover, being explicit and spelling out what cannot (or even should not in your opinion) be interpreted from the work might also be helpful. I think there is a space for one or two more sentence perhaps even in the Abstract as well.**

Connecting to the previous specification of methodology, we went through the manuscript and especially clarified our focus on first impressions of faces or photographs in multiple text sections. Especially the discussion benefitted from a more precise description of methodology in the sense of photograph perceptions. Therefore, the claim and limitations of our work is less vulnerable for politically motivated misinterpretations.

For instance, in the abstract it now reads (pg. 2):

*“It is important to note that these results should not be interpreted as any doubt about the profound experiences of discrimination and prejudices faced by minorities such as refugees. Instead, they underline the utility of an individual differences approach to better understand the circumstances under which devaluations of minoritized individuals suchs as refugees are amplified or reduced.”*

*In addition, on page 51, it now reads:*

*“However, we also highlight that the results should of course not be taken out of context to cast doubt on the actual numerous prejudice and discrimination experiences of minoritized individuals. Rather, this might have relevant implications for the development of effective interventions. As intergroup bias tends to be reduced through interpersonal contact (Islam & Hewstone, 1993; Pettigrew & Tropp, 2006; see also Paluck, Green, & Green, 2019) and confrontation (Czopp, Monteith, & Mark, 2006), it seems helpful to emphasize existing interindividual differences and idiosyncratic features within both majority and minoritized target groups.“*

For reasons of visibility, we also adapted the Conclusion, which now reads as following (pg. 60):

*“While we did not find a general devaluation of Middle Eastern refugees or Muslims, we do not claim in any way that this casts doubt about the numerous experiences of devaluation minoritized individuals face in daily life. Our results offer a limited perspective on perception outcomes based on face photographs under lab conditions leaving out crucial aspects of such daily discrimination. In addition, the presented studies found large variability in evaluations between both the evaluated targets and the evaluating perceivers. These interindividual differences were systematically related to targets’ cue characteristics and perceivers’ attitudes. Furthermore, refugee and nonrefugee targets who smiled more and were more physically attractive were evaluated more positively.”*

Furthermore, the references for first impressions on the basis of faces or photographs is now included in several text passages. For example, on pages 48, 49, 50, 51 and 57 it now includes a reference to face photographs:

*“[…] we conducted two studies to explore target and perceiver characteristics as predictors of first impressions of face photographs towards refugees.”*

*“In the present studies, we investigated perceptions and evaluations of refugee faces from photographs, that is […]”*

*“[…], the present study could not fully transfer these findings to a face perception setting of minoritized group individuals based on photographs.”*

*“In line with the idea that photograph perceptions of individual minoritized group members are not primarily driven by their group affiliations […]”*

*“[…] the perception of individual face photograph stimuli with ambiguous group affiliations is more ecologically valid for first impressions in daily life, which rarely include explicit labeling.”*

**E4: The paper is long. That might not be avoidable but I would just ask you to consider any last cuts to make this streamlined. I caught a few typos as well so check for those as well. For example, I think “they” is missing from Footnote #2.**

While we found and corrected several typos, such as the one kindly pointed out from Footnote #2, the length of the paper is difficult to substantially reduce. Due to the need for an extensive overview on two research streams with person perception and group-based research, together with rather complex methodology and analyses and the need to be very specific about the scope and limitations of the societal important topic, we were only able to slightly reduce the length. However, we cut some sections as a part repeating the hypotheses testing on page 48, a section on political debates on page 48 and one on face recognition on page 51. Therefore, we were able to cut three more pages.

**E5: I still missed having some grounding for interpreting effect sizes here (I do identify as a personality researcher, after all!). I understand the complexities with the modeling approach and lack of obvious standardization. Thus, this might be as good as it gets but just consider this issue in the final revision.**

We appreciate the remark on further ways to approach the interpretation of effect sizes. Accordingly, we again revised the section and added different ways to better place the presented effect sizes into context. The section was added in the general discussion of our second study and is also referenced in the first study’s general discussion.

On page 46f., it now reads:

“*Remarkably, the effect sizes we found indicate, as in Study 1, that individual-level target cues such as smiling and attractiveness played a much larger role in the evaluation of individual targets based on photographs than sheer group affiliation. The context and evaluation of effect sizes and their magnitude is a challenging topic for person perception studies including both cue-level and group-level characteristics. To better interpret the presented effect sizes, one can follow different approaches: First by looking at the variance ratios showing a substantial reduction of between-target variance about half when adding individual cues to the null model. Second, by calculating standardized regression coefficients using the presented (unstandardized) regression coefficients and the corresponding predictors' standard deviations12. Third, there are ways to define a region of practical equivalence (ROPE) around a point null hypothesis. This interval could then be compared to the confidence interval and help deciding about the null hypothesis. However, the definition of such a ROPE is dependent on the subject and scales of testing and therefore need to be informed by theoretical considerations or previous results. Due to the limited number of comparable studies and novelty in the comparison of cue-level and group-level characteristics, we are hesitant to define ROPEs, but would like to encourage future research to engage in the delineation of such ROPEs for different contexts and based on sufficient empirical studies.*”

The referenced footnote on the same page reads:

“*Because the presented criterion (positivity) is already standardized, one can calculate standardized regression coefficients (and their CIs) by multiplying the presented regression coefficients (and their CIs) with the corresponding predictor’s standard deviation (see the total sample section in Table 2 and Table 3).”*

We hope that you will now evaluate our manuscript as ready for publication in *Collabra:Psychology* and thank you very much for considering it.

Yours sincerely

Joscha Stecker (on behalf of all authors)

**Editor Final Decision—Accept**

April 8, 2021

Dear Joscha Stecker,

I have now had a chance to read over your response and the final version of your manuscript “First Impressions of Faces of Refugees are More Strongly Influenced by Target Cues and Perceiver Attitudes Than by Sheer Group Affiliation.” Thank you for your responsiveness to my remaining concerns. I am happy to write that your paper is now officially accepted for publication in Collabra: Psychology. Congratulations on this excellent work! I think it will make an important contribution to the literature and I look forward to seeing it published. I hope your experiences with Collabra: Psychology have been positive and that you will continue to consider it as an outlet for your work.

As there are no further reviewer revisions to make, you do not have to complete any tasks at this point. Our managing editor will contact you in case there are any pre-prodution file related questions. You will have an opportunity to check the page proofs before we publish your article. Thank you again for publishing in Collabra: Psychology.

Sincerely, Brent Donnellan