



HAL
open science

Accroître la généralisabilité de la psychologie en incluant les Africains

Adeyemi Adetula

► **To cite this version:**

Adeyemi Adetula. Accroître la généralisabilité de la psychologie en incluant les Africains. Psychologie. Université Grenoble Alpes [2020-..], 2024. Français. NNT : 2024GRALS014 . tel-04868248

HAL Id: tel-04868248

<https://theses.hal.science/tel-04868248v1>

Submitted on 6 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : ISCE - Ingénierie pour la Santé la Cognition et l'Environnement

Spécialité : PCN - Sciences cognitives, psychologie et neurocognition

Unité de recherche : Laboratoire interuniversitaire de Psychologie/Personnalité, Cognition, Changement social

Accroître la généralisabilité de la psychologie en incluant les Africains

Increasing the generalizability of psychology by including Africans

Présentée par :

Adeyemi ADETULA

Direction de thèse :

Annique SMEDING

PROFESSEURE DES UNIVERSITES, UNIVERSITE DE CHAMBERY

Directrice de thèse

Hans IJZERMAN

SCIENTIST, ANNECY BEHAVIORAL SCIENCE LAB

Codirecteur de thèse

Patrick FORSCHER

DOCTEUR EN SCIENCES COGNITIVES, UNIVERSITE GRENOBLE ALPES

Co-encadrant de thèse

Dana BASNIGHT-BROWN

United States International University-Africa

Co-encadrante de thèse

Rapporteurs :

OLIVIER KLEIN

FULL PROFESSOR, UNIVERSITE LIBRE DE BRUXELLES

ARMAND CHATARD

PROFESSEUR DES UNIVERSITES, UNIVERSITE DE POITIERS

Thèse soutenue publiquement le **24 juin 2024**, devant le jury composé de :

MAJA BECKER,

PROFESSEURE DES UNIVERSITES, UNIVERSITE TOULOUSE - JEAN JAURES

Présidente

ANNIQUE SMEDING,

PROFESSEURE DES UNIVERSITES, UNIVERSITE DE CHAMBERY

Directrice de thèse

OLIVIER KLEIN,

FULL PROFESSOR, UNIVERSITE LIBRE DE BRUXELLES

Rapporteur

ARMAND CHATARD,

PROFESSEUR DES UNIVERSITES, UNIVERSITE DE POITIERS

Rapporteur

CATHERINE BORTOLON,

MAITRESSE DE CONFERENCES HDR, UNIVERSITE GRENOBLE ALPES

Examinatrice

Invités :

JOY OWANGO

DIRECTRICE DE RECHERCHE, TRAINING CENTRE IN COMMUNICATION

HANS IJZERMAN

SCIENTIST, ANNECY BEHAVIORAL SCIENCE LAB



Acknowledgements

I would like to express my special gratitude to my parents Prof. and Deaconess Adetula for your care and prayers for me. Without you both, this dream can't be a reality.

My special appreciation to my beautiful and supportive sisters, Omotola, Oluwabusola, and Ibukun. Thanks for standing by me throughout my thesis program.

To my kids, Oreoluwa and Oluwafemi, it has been a long journey over the years and we have so much catching up to do. You guys made it easy for me to focus on my thesis. You are a blessing that I will forever be grateful for.

To Nwadiogo, I remember our discussion for me to move to France for my thesis. You started this journey by discussing with Hans on my behalf and later made the introduction and today your kind gesture and efforts have borne the desired fruit. Thank you so much for showing me around Grenoble and for the friendly discussion and advice about life and career.

Thank you, Tetiana, for always checking on me and for the coffee breaks to make sure I'm not feeling lonely.

My special gratitude to Juliet, your prayers and love strengthen me.

To my supervisors, Hans, Patrick, and Dana, and the one supervisor I didn't see coming Annique, thank you all for your guidance, kindness, and support over the past 4 years.

Hans, what a journey! I remember our first call to discuss a potential thesis. Despite having ideas about social thermoregulation, you insisted that the best course for me and where you see greater career potential is replication. Although I reluctantly agreed to pursue a replication-focused thesis, today, that decision proved to be one of the best decisions for my career. I can't forget in a hurry how you stood by me when my grant was a mess. You kept encouraging me and providing resources and guidance to scale through this program. I know working with me was demanding, your patience, willingness to teach and even sometimes "compromise" allowed us to complete this project. Including yourself, you assembled a super supervision team any doctoral student could wish for. I'm grateful and look forward to other projects with you.

Patrick, my friend, and supervisor, as you promised, you made sure I finished the thesis. In the early days of my thesis, despite my confusion about the scale of the project, you always calmly reassured me that it was doable. You stood by me all through the project providing assistance, resources, and guidance. You know we have just begun the research journey with so many projects to collaborate on. Thank you for this one!

With Patrick and Hans, sometimes it could be stormy deciding what to do, the writing, and rigor (I'm looking forward to meeting my supervisees) but for Dana, who is always on my side and ensures the confusion and my inadequacies don't trouble me. Thank you, Dana, for your guidance, support, and helpful feedback.

To Annique, you stepped in at a crucial time, and from then on you showed unwavering support and commitment to my thesis. I would have loved to have you from the start but I believe we have more to accomplish together in the future. Thank you so much!

To my Comité de Suivi Individuel committee, Maja, and Arnaud, just like yesterday when I requested you to be part of my Comité de Suivi Individuel team, you both without hesitation accepted to be part of my thesis. Thank you for your feedback even at the shortest notification, you provided useful evaluations and ensured the project went according to plan.

To my CORE Lab people, Nathan, Olivier, Alessandro, Miguel, Maria, Tetiana, Mae, Ronan, Patrick, Hans, and others, I may have missed thank you for the memorable moments shared.

To all my friends and colleagues at LIP/PC2S, it was a great pleasure knowing and working with you all. Some of you provided beautiful drawings of my face. Special thanks to Anne, Dominique, and Céline for your support over the years.

Thanks, Orélie and Chris, for the excellent work. Yes, I know, EDISCE is responsible for doctoral students but your kindness and responsiveness at my shortest notice was very much appreciated.

I want to recognize one of my thesis project partners, the Collaborative Replication and Education Project (CREP), and Dr Jordan Wagge, the executive director of CREP. Jordan you were instrumental to the success of the CREP Africa study. Thank you for your support, your time and resources, and your contributions to the project.

I'm grateful to all donors and patrons via my GoFundMe and Patreon. Without your financial support, there is no way I can complete this thesis. Thanks for allowing humanity to prevail.

Special thanks to Dr. Thuy-vu Nguyen, for your support. It means a lot to me.

To all our CREP Africa and ManyLabs Africa studies collaborators, Thank you!

Abstract

Most psychological claims were originally discovered in the US and Europe. The extent to which these findings generalize to Africans is unclear. One initiative that is part of psychology's credibility revolution, big team science, attempts to improve generalizability through large, geographically distributed collaborations. However, African participation remains poor in these initiatives. The objectives of this thesis are to fix this problem in big team science by 1) building African researchers' capacity, 2) improving African participation, and 3) centering African thoughts and priorities.

To accomplish these objectives, we conducted two big team science replications. First, we partnered with the Collaborative Replication and Education Project (CREP), an open science-driven replication platform designed to provide training to users, to 1) familiarize our African collaborators with the CREP model and 2) engage them in an Africa-wide replication of the "mere-trace hypothesis" that moral judgment of harm transgressions are more sensitive to the violation severity than purity transgressions. Twelve contributing African sites from Burkina Faso, Kenya, Morocco, Nigeria, and Tanzania sampled 783 participants after exclusions. We found that, consistent with the original claim using US samples, our African participants judged severe harm transgressions as more wrong than less severe ones, but were not as sensitive to severity for purity transgressions.

The CREP study enabled us to build African collaborators' capacity, preparing us for a second study, ManyLabs Africa. Guided by consultations and preset criteria, we selected 18 sexual behavior claims originally discovered and of priority to our African collaborators. We sampled 5806 participants after exclusions across Africa, Europe, and North America. Only 4 of 18 claims replicated successfully across regions: 1) people predicted that parents should not force

their teenage daughter to abort a pregnancy, 2) male college students aged 18 - 20 reported more positive attitude towards condom use than females, 3) women predicted that they would be relatively 3) willing to forgive their husbands for a emotionally uninvolved affair, and 4) willing to forgive a husband expressing regret for the affair. We found a relatively low amount of region-specific variability between sites for these claims, though we could only sample a few countries and our participants were mostly educated, young, and urban.

Big team science can complement the ongoing capacity-building efforts in Africa and expand African theories but there is a need to improve research facilities in Africa. African researchers can leverage the open infrastructure and resources and extensive network of labs and collaborators to advance their capacity and careers. We drew on this project experience to recommend resources, training, research tools, and ethics interventions to remove research barriers for large-scale replication and to foster inclusion.

Keywords: Africa, Generalizability, Psychology, Multi-site replications

Résumé

La plupart des affirmations psychologiques ont été découvertes aux États-Unis et en Europe. La mesure dans laquelle ces résultats s'appliquent aux Africains n'est pas claire. Une initiative qui fait partie de la révolution de la crédibilité de la psychologie, la science des grandes équipes, tente d'améliorer la généralisabilité par le biais de vastes collaborations réparties géographiquement. Cependant, la participation africaine reste faible dans ces initiatives. Les objectifs de cette thèse sont de résoudre ce problème en 1) renforçant les capacités des chercheurs africains, 2) améliorant la participation africaine, et 3) centrant les pensées et les priorités africaines.

Pour atteindre ces objectifs, nous avons mené deux études de réplication par le biais de la science des grandes équipes. Tout d'abord, nous nous sommes associés au Collaborative Replication and Education Project (CREP), une plateforme de réplication scientifique ouverte conçue pour former les utilisateurs, afin 1) de familiariser nos collaborateurs africains avec le modèle CREP et 2) de les engager dans une réplication à l'échelle de l'Afrique de l'hypothèse de la « simple trace » selon laquelle le jugement moral des transgressions de dommages est plus sensible à la gravité de la violation qu'à celle des transgressions de pureté. Douze sites africains du Burkina Faso, du Kenya, du Maroc, du Nigeria et de la Tanzanie ont échantillonné 783 participants après exclusion. Nous avons constaté que, conformément à l'affirmation originale utilisant des échantillons américains, nos participants africains jugeaient les transgressions de dommages graves comme plus mauvaises que les transgressions moins graves, mais n'étaient pas aussi sensibles à la gravité des transgressions de pureté.

L'étude CREP nous a permis de renforcer les capacités de nos collaborateurs africains, nous préparant ainsi à une seconde étude, ManyLabs Africa. Guidés par des consultations et des

critères préétablis, nous avons sélectionné 18 allégations de comportement sexuel découvertes à l'origine et jugées prioritaires par nos collaborateurs africains. Après exclusion, nous avons échantillonné 5806 participants en Afrique, en Europe et en Amérique du Nord. Seules 4 des 18 allégations ont été reproduites avec succès dans toutes les régions : 1) les gens ont prédit que les parents ne devraient pas forcer leur fille adolescente à avorter, 2) les étudiants masculins âgés de 18 à 20 ans ont signalé une attitude plus positive à l'égard de l'utilisation du préservatif que les femmes, 3) les femmes ont prédit qu'elles seraient relativement 3) disposées à pardonner à leur mari pour une liaison sans implication émotionnelle, et 4) disposées à pardonner à un mari exprimant des regrets pour la liaison. Nous avons constaté une variabilité régionale relativement faible entre les sites pour ces affirmations, bien que nous n'ayons pu échantillonner que quelques pays et que nos participants aient été pour la plupart instruits, jeunes et urbains.

La science des grandes équipes peut compléter les efforts de renforcement des capacités en cours en Afrique et développer les théories africaines, mais il est nécessaire d'améliorer les installations de recherche en Afrique. Les chercheurs africains peuvent tirer parti de l'infrastructure et des ressources ouvertes ainsi que du vaste réseau de laboratoires et de collaborateurs pour faire progresser leurs capacités et leur carrière. Nous nous sommes appuyés sur l'expérience de ce projet pour recommander des ressources, des formations, des outils de recherche et des interventions en matière d'éthique afin d'éliminer les obstacles à la recherche en vue d'une réplique à grande échelle et de favoriser l'inclusion.

Mots-clés : Afrique, Généralisabilité, Psychologie, Répliques multi-sites

Table of contents

	Pages
<i>Acknowledgements</i>	<i>i</i>
<i>Abstract</i>	<i>iv</i>
<i>Résumé</i>	<i>vi</i>
<i>Table of contents</i>	<i>viii</i>
<i>List of Tables and Figures</i>	<i>xii</i>
<i>Chapter 1: Introduction - Why we should investigate African claims</i>	1
1. How psychology can benefit from African insights.....	3
2. How to nurture African psychology	5
3. Replication of African-origin effects	6
4. Sustaining African participation	7
<i>Chapter 2a: CREP Africa study preparatory work</i>	8
1. The Collaborative Replication and Education Project Model.....	8
2. Pilot Study 1: Building Research Infrastructure.....	10
2.1 Recruitment of Collaborators.....	11
2.2 Creation of Training Materials.....	11
2.3 Infrastructure and Resources Assessments.....	12
3. Pilot Study 2: Study Selection	15
3.1 Study Selection	15
3.2 Re-analysis of Rottman and Young (2019)	17
3.3 Power Analysis	19
<i>Chapter 2b - CREP Africa study: A paradigmatic replication of Rottman and Young (2019)</i>	21
1. Method	23
1.1 Participants	23
1.2 Materials.....	24
1.3 Translation and Cultural Adaptation of Materials	25
1.4 Procedure	26
2. Results.....	29
2.1 Analysis plan.....	29
2.2 Sample and sample demographics	30
2.3 Domain categorizations, and domain and dose analyses: Manipulation checks	32

2.4 Main and interaction effects analyses: Confirming the effect of domain, dose, and domain by dose interaction on moral wrongness judgment	33
2.5 Determining if we replicated Rottman and Young	35
2.6 By-participant, by-item, and by-site random effects	35
2.7 Exploratory analyses: What are important moral issues to Africans	36
3. Discussion	37
3.1 Study findings	37
3.2 Data quality	38
3.3 Theoretical implications.....	39
4. Recommendations	40
5. Conclusion	41
<i>Chapter 3a: Determining an approach for the ManyLabs Africa study: A Global South perspective.....</i>	<i>42</i>
1. Summary	42
2. Introduction	43
3. The consultants.....	44
4. Method and procedure: Conducting the consultations	45
5. Foci of the interviews and general conclusions	45
5.1 Individual consultants' general opinions.....	46
5.2 Studies to focus on.	46
5.3 Preferred study selection approach.....	47
6. Conclusion.....	49
<i>Chapter 3b - ManyLabs Africa study nomination, evaluation, and assessment report</i>	<i>50</i>
1. Step 1: Study nomination.....	52
1.1 Respondents	52
1.2 Response analysis	53
1.3 Exclusion and ranking of nominated research topics	53
1.4 Study search challenges.....	55
2. Step 2: Evaluation of nominated studies/effects	56
3. Step 3: Study assessment	57
3.1 Respondents	58
3.2 Method and analysis	58
4. Results.....	59
4.1 Overall conclusion: Study selection decision	59
5. Study selection decision summary	66
6. Conclusion.....	68
<i>Chapter 4: ManyLabs Africa study: Investigating the replicability and generalizability of African effects.....</i>	<i>70</i>
1. The problem with generalizing from North America (and Europe)	71

2. Selected studies for our replication 73

3. Replication Hypotheses..... 77

4. Method 82
 4.1 Participants, Power analysis, Recruitment, and Ethics..... 82

5. Procedure..... 88

6. Measures 89
 6.1 Teenage pregnancy study measures..... 89
 6.2 Condom use negotiation study measure..... 90
 6.3 Extramarital sex study measure 91
 6.4 General measures..... 91

7. Cultural adaptation and translation of our measures 92

8. Data analysis plan 93

9. Results..... 96
 9.1 Teenage pregnancy study results 96
 9.2 Condom use negotiation study results 104
 9.3 Extramarital sexual relationship study results..... 111
 9.4 Auxiliary analyses results..... 115

10. Discussion, Recommendation, and Conclusion 116
 10.1 Discussion 116
 10.2 Limitations..... 123
 10.3 Recommendations..... 124
 10.4 Conclusion..... 124

Chapter 5: Synergy Between the Credibility Revolution and Human Development in Africa..... 126

1. The research landscape in Africa..... 128

2. The credibility revolution in North America and Europe..... 131

3. How open science practices can contribute to human development 133

4. Broadening the open science tent in Africa..... 139

5. Conclusion..... 146

Chapter 6: General Discussion, Conclusion and Recommendations 147

1. Need for more scientists and capacity building 147

2. Need to improve African research accessibility and visibility 148

3. Need to engage Africans for improved participation 149

4. Need for more infrastructure and resources for improved productivity 150

5. Need to adopt open science tools and practices for credible research..... 151

5. General Conclusion: Contributions 152
 5.1 Theoretical contributions 157
 5.2 Methodological contributions..... 158
 5.3 Capacity building contributions..... 159

5.4 Research policy contributions 161

6. The future of African big team science..... 162

References 164

Appendix I..... 186

1. Consultants' general opinions of the proposed ManyLabs Africa study..... 186

Appendix II 189

1. Consultants take on studies to focus on..... 189

Appendix III..... 191

1. Consultants take on a preferred study selection approach 191

Appendix IV 193

Appendix V..... 206

Appendix VI..... 212

1. Teenage pregnancy study results for full sample..... 212

1.1 Agreement proportions for the full sample. 212

1.2 Region comparisons for agreement proportions for full sample 215

Appendix VII..... 218

1. Condom use negotiation study results for the full sample 218

1.1 Main effect of gender 218

1.2 Gender by region comparison for the entire participants..... 219

Appendix VIII 225

1. Extramarital sexual relationship study results..... 225

1.1 Main effect of emotional involvement and regret expressed on forgiveness for the full sample 225

List of Tables and Figures

Tables	Pages
Table 1 Big-team science initiatives in psychology	2
Table 2 Re-analysis of Study 1, Study 2 and Study 3 from Rottman and Young (2019).....	18
Table 3 A set of 32 scenarios for harm and purity violations.....	24
Table 4 Domain by dose coefficient replication interpretation	30
Table 5 Number of participants per site, their gender distributions, and their average age	31
Table 6 Linear Mixed Effects Model results for participants’ wrongness judgments and the original study meta-analysis aggregate.....	34
Table 7 Selected study assessment parameters’ means and SD for overall and per continent	63
Table 8 Mgbokwere, Esienumoh, and Uyana (2015) study’s frequencies, percentages, and our calculated percentage CI.....	74
Table 9 Replication hypotheses	79
Table 10 Original and replication samples description.....	85
Table 11 Number of participants per country	86
Table 12 Participants’ demographics.....	86
Table 13 Replication study statistical analyses deviations	95
Table 14 Teenage pregnancy replication results for parents aged 35 and above, and the replication decision	98
Table 15 Teenage pregnancy replication region comparisons for parents aged 35 and above, and heterogeneity test	102
Table 16 Condom use negotiation results for the subgroup analyses, and the replication test	106
Table 17 Extramarital sexual relationship focal effect results for heterosexual women only	112
Table 18 Five categories of open science practice and initiatives that fit each category.	133
Table 19 Barriers to expanded research capacity and interventions to address those barriers	139
Table 20 Respondents’ study adaptability comments by continent for the extramarital affair study.....	193
Table 21 Respondents’ study adaptability comments by continent for the teenage pregnancy study.....	195
Table 22 Respondents’ study adaptability comments by continent for the civic engagement study.....	198
Table 23 Respondents’ study adaptability comments by continent for the self-disclosure study	201
Table 24 Respondents’ study adaptability comments by continent for the condom use negotiation study.....	203
Table 25 Respondent’s additional comments on the extramarital affair study by continent.	206
Table 26 Respondent's additional comments on the teenage pregnancy study by continent	207

Table 27 Respondent's additional comments on the civic engagement study by continent ..208
 Table 28 Respondent's additional comments on the self-disclosure study by continent209
 Table 29 Respondent's additional comments on the condom use negotiation study by continent.....210
 Table 30 Teenage pregnancy replication results for the full sample213
 Table 31 Teenage pregnancy replication region comparisons for the full sample, and heterogeneity test216
 Table 32 Condom use negotiation results for the overall analyses and the replication test ..220
 Table 33 Extramarital sexual relationship focal effect results for the entire participants227

Figures	Pages
Figure 1 Power curves in three different scenarios.....	20
Figure 2 Violin plot of wrongness rating as a function of domain and dosage for our replication and Rottman and Young (2019) predicted means	34
Figure 3 Selected study assessment factors per continent and for all regions	62
Figure 4 Willingness to forgive mean ratings as an operation of the three factors (sourced from Vera Cruz, 2018).....	77
Figure 5 Attitude to condom use by gender per region for adolescent students.....	108
Figure 6 Perceived control over condom use by gender per region for adolescent students.	108
Figure 7 Self-reported condom-protected sex behavior by gender per region for adolescent students	109
Figure 8 Willingness to forgive mean ratings as an operation of regret expressed and emotional involvement and for region comparison	114
Figure 9 Attitude to condom use by gender per region for the full sample	222
Figure 10 Perceived control over condom use by gender per region for the full sample	223
Figure 11 Self-reported condom-protected sex behavior by gender per region the full sample	224
Figure 12 Willingness to forgive mean ratings as an operation of regret expressed and emotional involvement and for region comparison for full sample.....	229

Chapter 1: Introduction - Why we should investigate African claims

Psychology focuses on understanding human behavior and has a moral obligation to study all humans (Rad et al., 2018). However, psychology has consistently focused on and is dominated by questions from a 4% population — the United States and, to a lesser extent, Western Europe (Thalmayer et al., 2021; IJzerman et al., 2021). Of note is the African continent which houses 17% of the global population, yet is consistently neglected in psychology (Singh et al., 2023). For example, in one audit of top psychology journals, less than 1% of first authors and research samples were from Africa (Thalmayer et al., 2021). When psychology researchers do focus on Africa, they too often take one particular, and perhaps problematic approach: they take a claim developed with one population and test whether that claim generalizes to Africa. This original population is usually from North America (and to a lesser extent Europe; IJzerman et al., 2021) and often college students (Thalmayer et al., 2021; Henrich et al., 2010).

As further evidence that this happens, consider a style of project known as “big-team science”. These are highly collaborative projects that pool together resources to achieve samples of exceptional national diversity (Forscher et al., 2022). Big-team science projects are an interesting case because they often have a cross-cultural focus and make a conscious effort to recruit international samples with a few African samples. However, as shown in Table 1, this style of project almost universally investigates whether effects discovered in North America or Europe generalize to other populations. Almost always, the lead investigators are also from North America or Europe.

Table 1 Big-team science initiatives in psychology

Initiatives	N of studies	Origin of Effects		Project leads	
		N/A & E	Africa	N/A & E	Africa
Collaborative Replication and Education Project, https://osf.io/wfc6u/	17	17	0	9	1
ManyBabies, https://manybabies.github.io/	17	14	0	31	0
ManyLabs (https://osf.io/wx7ck/ , https://osf.io/8cd4r/ , https://osf.io/ct89g/ , https://osf.io/8ccnw/ , https://osf.io/7a6rd/)	5	62	0	3	0
Psychological Science Accelerator, https://psysciacc.org/	9	11	0	18	0

Note. N/A & E- North America and Europe
 The origin of the effects and research leads for four big-team science initiatives in psychology are almost entirely from North America and Europe.

This style of research can legitimately advance science. However, it also runs the risk of creating “unknown unknowns”: gaps in knowledge that we aren’t even aware of because we haven’t bothered to take seriously the idiosyncrasies and patterns of thought that are unique to Africa. It also risks recreating a harmful dynamic held over from colonialism, in which African labor goes toward generating knowledge and advances that are mostly useful and relevant to richer white people.

The current thesis broadly aims to improve generalizability of psychology. Specifically, the objectives are to 1) improve African collaborators' capacity in open science practices and big team science, 2) test the applicability and heterogeneity of psychology claims within and outside Africa, and 3) recommend interventions to facilitate open science adoption and African participation in psychology. These objectives were accomplished in 6 chapters. In Chapter 1,

we drew attention to the problem of African underrepresentation, the generalizability problem, and why and how the current big team science replication trends in psychological science are Western-focused. In Chapter 2, we conducted an Africa-wide replication via an open science-driven model to familiarize our African researchers with some open science practices and build their research capacity. By doing so, we can help address African underrepresentation, low research capacity, and potential lack of generalizability of findings. However, the Africa-wide study is typical of Global North to South generalization which often leads to biases in questions being investigated and misplaced priority for Africans. To provide test generalizability from an African population to a Western one, in Chapter 3, we first determined an approach for selecting African-preferred studies for replication. To do so, we consulted researchers working in the Global South to determine a study focus and developed a guide/procedure for our African study selection. This approach also helps to identify African priorities and address the problem of the “unknown unknowns”. In Chapter 4, we then investigated the replicability and generalizability of three claims previously investigated among Africans (that were selected following our guide) to other populations. In Chapter 5, owing to our investigations and experience, we discussed the need for research development in Africa. Lastly, In Chapter 6, we reflected generally on four takeaways, whether we accomplished our set goals and the thesis contributions to research.

1. How psychology can benefit from African insights

Focusing on the African experience is the moral and ethical thing to do. However, this focus could also benefit psychology as a whole. African cultures are culturally rich and complex, and developing theories to explain the patterns of thought that arise from these cultures could lead to theories that would not easily arise from North American and European

minds. Take *Ubuntu*, a South African concept of moral and communal obligation to others, or “the belief in a universal bond of sharing that connects all humanity”, with a self much more permeable than in North America and Europe, as “I am” is not a rigid subject, but a dynamic self-constitution dependent on this otherness creation of relation and distance” (Eze, 2010). Psychological theories developed to explain Ubuntu may require a different sense of moral value than is usually assumed in theories from North America and Europe. As another example, the Yoruba personhood concept of *Omoluabi* (Adeniji-Neill, 2011) describes character, courage, and hard work. An Omoluabi is a person who respects honor, works hard, and gives to the community. Both Ubuntu and Omoluabi may perhaps have their equivalents in North America and Europe, but without proper study, we don't know. They are also directly relevant to the African experience, unlike some current research in big-team science. Had African researchers selected topics relevant to "morality" to examine on a global scale, they may not have selected "trolley dilemmas" that have questionable relevance to African moral reasoning (Bago et al., 2022).

Beyond research question diversity, psychology can benefit from the growing population of Africa to nurture good research psychologists. Africa is estimated to be 26% of the world population by 2050 (United Nations Department of Economic and Social Affairs, Population Division, 2022). Moreover, psychology can compete with growing substitute disciplines such as public health and development economics by involving more Africans in global research (Forscher et al., 2021). As building research capacity was one of the present thesis's objectives, our Africa-wide CREP study and the provided training resources were to familiarize and build the capacity of our participating African researchers in adopting open science practices (see Chapters 2a and 2b).

2. How to nurture African psychology

Generalizing from Africa to North America and Europe requires a robust psychology research literature in Africa from which to generalize. Some of this literature already exists; the International African Journal maintains repositories from 28 African countries (<https://www.internationalafricaninstitute.org/repositories>) and the African Journals Online maintains a list of 293 open access African journals (<https://www.ajol.info/index.php/ajol>).

However, existing African psychology may look different from the typical scientific psychology study in North America and Europe. For example, some African psychology focuses on more applied topics and more often uses qualitative approaches. The applied focus means a greater emphasis on subfields like clinical, forensic, and occupational psychology. Some of this research does not even call itself “psychology.” Despite investigating psychological topics, it brands itself as behavioral economics, development economics, and public health.

Material realities also constrain what is easily possible in African research institutes and with African populations. These include inadequate infrastructure, communication problems due to Africa’s incredible language diversity, a lack of access to research literature, intermittent Internet, a lack of external funding leading to self-funded research, decontextualized measures developed in North America and Europe, and modest to low literacy (Onie, 2020). Finally, many African countries are held back by historical and political events that have starved them of funding and infrastructure, leading to sporadic growth of African psychological science (Adetula et al., 2021). In Chapter 2a, we detailed our efforts to work within constraints.

3. Replication of African-origin effects

There is, however, one step to take immediately to foreground African psychology: instead of always generalizing from North America and Europe to Africa, why not try generalizing from Africa to North America and Europe? This might involve selecting effects that support important claims in African research and testing their replicability in North America and Europe. Replicating African-originated claims means identifying what questions are paramount to Africans as well as understanding African thought processes and methods that could help discover the “unknown unknowns” that exist in African psychology. Hence, a first pragmatic step will be to consult with researchers from the Global South on research interest in this world region and how to navigate a study selection (see Chapter 3a). Based on our consultation, we asked African researchers to suggest research areas and nominate studies that originated in African countries (see Chapter 3b). This served as the basis of the selection of effects that will serve as targets of replication, both in Africa and in North America and Europe. Our nomination survey revealed that our African colleagues were interested in both issues that feature prominently in psychology in North America and Europe, such as sexual relationship and behavior, gender, parenthood, crime and criminality, and family, but also in issues that do not feature prominently, such as hatred and disharmony, land use conflicts, polygamy, corruption, poverty, and spirituality. This nomination process reveals some of the important topics that, although they affect the lives of many, are neglected by psychological science from just a small part of the world. More telling is the dichotomy we observed in study to replicate preferences between Africans and the Europeans/North Americans when selecting our final studies (see Chapter 3b). While a teenage pregnancy study was the most preferred for Africans, a self-disclosure study was the most preferred by Europeans and North Americans. In Chapter 4, we discussed our methods and findings on 3 African-originated claims replication.

4. Sustaining African participation

Our approach is only a first step toward providing better representation of African researchers, concepts, and people in psychology. Succeeding in the grand goal of making psychology more representative of Africa will require much more, including providing resources, training and research tools, and upholding ethical practices. We elaborate on these steps further in Chapter 5.

Chapter 2a: CREP Africa study preparatory work

One potential reason for the lack of African researchers and participants is that psychology research is “nascent” in most African countries (Nsamenang, 2007), a state of affairs that reflects a continent starved of the resources to compete with its contemporaries (UNESCO, 2021). This nascent status means that little capacity to conduct psychology research in Africa is available (Adetula et al., 2021). A first step to address this challenge is to conduct research in which the methods are well developed, as is the case of replication studies while involving researchers from the target populations. For this part of the project, we first conducted some preparatory work that includes knowing the resource and infrastructure realities of our African collaborators, creating training materials and familiarizing African collaborators with our tools-of-choice for the replication study, as well as developing a procedure for selecting a study that is feasible, reproducible, and applicable for a replication study in Africa. We then conducted a paradigmatic replication (Vohs et al., 2021) of studies of purity- and harm-based moral judgments from Rottman and Young (2019) in five African countries. We used a training model developed by the Collaborative Replications and Education Project (CREP), a platform that emphasizes “Big Team” replication studies to teach cutting-edge research practices (Wagge, Baciu et al., 2019; Wagge, Brandt et al., 2019) to test the generalizability to Africans on a set of findings on moral reasoning about harm and purity transgressions.

1. The Collaborative Replication and Education Project Model

This project used a model developed by the Collaborative Replication and Education Project (CREP). We chose the CREP model because it is a highly structured model for training

in cutting-edge research methods and gives users many opportunities to provide feedback to improve the model. The latter is particularly useful when deployed in new environments.

As it was originally conceived, the CREP model assumes little to no prior experience in either open science practices or scientifically-oriented, quantitative psychology (see Wagge, Brandt, et al., 2019). CREP trainees learn by doing, both open science practices and scientific psychology by conducting a replication study. The trainees organize their project and its materials via a page on the Open Science Framework (OSF). In the first stage of organizing their project, trainees create a copy of the general CREP project by “forking” it. The trainees organize their study materials by filling out each component in the forked project. This includes the full study materials, including any surveys, protocols, and pictures of the materials, a videotape of an experimenter completing the procedure, ethics approval, and a written preregistration. The project is then reviewed for accuracy by an instructor and submits the page to the CREP board. The CREP board conducts a second set of reviews and gives the trainees feedback about the structure and accuracy of the OSF page.

After the pre-data-collection review process is complete, the trainees pre-register the study by creating a frozen copy of the complete OSF project and then proceed to data collection. Under the traditional CREP model, each trainee recruits at least 100 participants, whose responses are stored in a dataset that they themselves manage. Each trainee documents their dataset with a codebook, and creates their own data analysis script. These materials are uploaded to their project pages, which are submitted to the CREP board for another round of review. Once the post-data-collection reviews are complete, the trainees receive a certificate of completion from the CREP board. The full CREP project consists of a meta-analytic synthesis of the results reported by each trainee.

We adapted this model to help build research capacity among our collaborating researchers. Instead of meta-analyzing the effect sizes, we decided to manage a single dataset

centrally, both to allow for the use of Linear Mixed Effects Models on the central dataset in an effort to minimize errors in the merging of the dataset and in the analysis pipeline. To gain the benefits of the CREP training, our collaborators completed the other elements of the CREP procedure. We provided materials - like methods, instruments, or training materials - needed for the present study. To ensure the studies were feasible, we decided that researchers could participate if they had basic infrastructure and required resources such as access to a lab and research facilities, space for in-person survey, reliable internet service, and internet connecting devices.

Opportunities for training on how to do replications in quantitative research psychology is not common in all African countries. Thus, we needed to do some preparatory work to help build the capacity to do this kind of research and to help better understand the needs or the current capacity. We report here our activities to build this capacity, reporting this preparatory work as “pilot studies”. Below, we describe our methods for building research capacity among our collaborators (Pilot Study 1), and our method for selecting a study to replicate (Pilot Study 2).

2. Pilot Study 1: Building Research Infrastructure

While many researchers have focused on improving the reproducibility of our science, only few have focused on the generalizability of our science. We suspect this may be the case because collecting data in hard-to-reach samples is challenging and that we first need to create common ground with the researchers we work with. We think that the CREP model provides a reasonable opportunity to challenge open science-related skills to those researchers previously unfamiliar with the skills, while it also allows collaborators that may have other skills to further improve the CREP model. Despite these advantages, training researchers in this model requires additional work and interaction, which we describe here.

2.1 Recruitment of Collaborators

To recruit collaborators for our project, we first reached out to international initiatives like the Psychological Science Accelerator (Moshontz et al., 2018) and ManyLabs studies (Ebersole et al., 2016; Klein et al., 2014, 2018, 2022). We then reached out to psychology associations in Africa to help distribute our research collaboration invitation to members. Lastly, we contacted African researchers in our network directly. We sent out emails to all with a sign-up form (<https://osf.io/bhjnX/>) for persons interested in the project. Out of the 70 African researchers who signed up for the project, 56 collaborators from 23 sites from 11 countries (Burkina Faso, Kenya, Malawi, Morocco, Mozambique, Nigeria, République démocratique du Congo, Rwanda, Tanzania, Togo, and South Africa) participated in the project.

2.2 Creation of Training Materials

We created training materials and organized hackathons and workshops on the CREP and open science practices for our African collaborators. We first organized a series of hackathons and workshops where some of the collaborators from various African countries participated in open science discussions and provided input at different stages of creating training materials. These training materials include training videos fashioned for African researchers and subtitled in seven African languages (Arabic, Chichewa, Hausa, Igbo, Portuguese, Swahili, Yoruba), syllabi, and a step-by-step CREP guide (<https://osf.io/wrntg>). The training materials are accessible and downloadable from our OSF Tutorial Videos page: (<https://osf.io/8akz5/>) and from YouTube (https://youtu.be/_x8dZxLmc-o, <https://youtu.be/QLzDkHwepVY>, <https://youtu.be/QmwmEPAvzcl>, and <https://youtu.be/9jSLuNEUzjE>). We also created resource syllabi for our collaborators' courses with their input (<https://osf.io/rfudt/>).

2.3 Infrastructure and Resources Assessments

To ensure that collaborators can realistically conduct these studies as well as to guide critical decisions throughout the project, we carefully consider the needs and potential barriers noted during our interactions during hackathons and workshops with our African collaborators. The central question is whether participating labs or researchers have the needed facilities and resources to be able to successfully conduct the study locally. The participating researchers are regarded as “Sites” responsible for coordinating all the research activities locally including to set up the Study’s OSF CREP page for the sites, documentations, translation and adaptation of instruments, data collection, and so forth. We estimated that we required a minimum of 10 sites to attain proper power to test the focal effects.

We surveyed our current 57 collaborators from 12 African countries (Algeria, Cameroun, Egypt, Ethiopia, Ghana, Kenya, Malawi, Morocco, Nigeria, Rwanda, Tanzania, and South Africa) for infrastructure and resource availability in these sites. Of these 57 collaborators, 23 from Egypt, Malawi, Morocco, Nigeria, Tanzania, and South Africa responded to our survey. Overall, our findings were encouraging and showed that the present replication study is very feasible. We discussed these needs, how we have or would address these needs, and what limitations there are under the following categories: resource availability, online study feasibility, translation and translation competence, familiarity with open science practices, and IRB / Ethical Approval Requirement.

2.3.1 Resource Availability.

Surveyed collaborators estimated an average expense of \$653 USD for participant compensation, internet access, research assistants, IRB review fees, paying translators, and so forth. This is important, as this research project is not funded, therefore we relied on participating labs and/or contracted collaborators to fund the project locally. We restricted 75%

of participating collaborators to collaborators contracted to research labs where some of the required resources are available. 8 out of 23 (35%) respondents have contracts of which only one had funds available for research projects. As a result, we planned to recruit more collaborators contracted to research labs.

In terms of whether these labs are equipped to conduct the study, the 8 collaborators contracted to labs reported having a designated room to conduct in-person surveys, and 6/8 have contracted research assistants in their labs. 6/8 reported devices connecting the internet (such as a computer), and 5 of them said there is a computer available for online study. The internet quality could constitute barriers as only 2 out of the 8 collaborators reported that the internet service is sufficiently reliable. However, 14 out of the 23 respondents received internet access subscription grants of \$100 USD and \$50 USD from The Society for Personality and Social Psychology (SPSP; 12 recipients) and The Society for the Improvement of Psychological Science (2 recipients) respectively.

2.3.2 Online Study Feasibility.

Online study is reported to be prevalent in our collaborators' localities with 96% of the respondents indicating at least a moderate feasibility and 91% stating that the local populations have a moderate willingness to participate in online study. However, in designing the survey, we enabled the smartphone-friendly option for Qualtrics surveys in order to accommodate the very large percentage of respondents (83%) who reported the use of smartphones for online surveys. Although 78% collaborators indicated the need to compensate participants, we were unable to pay participants compensation because there were no funds, and payment across national boundaries is relatively administratively difficult, and we did not have the capacity to make this work. Collaborators reported adopting various methods and procedures to recruit participants for online and in-person studies. None of the participant recruitment methods and

procedures mentioned should have implications for the study. However, CREP reviewed the data collection procedures before the actual data collection.

2.3.3 Translation and Translation Competence.

We recruited 15 collaborators with a C1 or C2 proficiency level (Council of Europe, 2020) in the target language and reported to be experienced or very experienced in research materials translation to help translate the questionnaire in the following local African languages: Arabic, Chichewa, French, Swahili, and Portuguese.

2.3.4 Familiarity with open science practices.

For this current project, participating collaborators familiarized themselves with specific open science practices like OSF, preregistration, Registered Report, and research collaboration. Only 2 out of 23 respondents have engaged in either preregistration or Registered Report. Seven out of the 23 respondents have engaged in open sourcing. 13 out of 23 (57%) have participated in some multisite studies either to coordinate lab/site or collect data or translate research materials. Though the report of previous participation in crowdsourced multisite studies is encouraging, the report indicates more training is required. Although 12 of the 23 respondents participated in our SPSP hackathon and workshop on open science and CREP, we made available training resources and also the research lead was available for consultation.

2.3.5 IRB/ Ethical Approval Requirement.

While collaborators are either not sure or have no such requirements for ethics approval, collaborators at the following labs/departments indicated the need to obtain ethical approval: Kamuzu University of Health Sciences (formerly college of medicine, University of Malawi), The Open University of Tanzania, Behavioral and Economic Decision-Making Lab (BEDMLab), The American University In Cairo, Alex Ekwueme Federal University Ndufu

Alike, and Alex Ekwueme Federal University Teaching Hospital, Abakaliki. There are university/college owned IRB review institutes at these sites. Working with collaborators at these sites, Dr. Basnight-Brown, one of the project leads, led the ethics consideration and IRB application. It took an average of 5 weeks to obtain approval from these institutes. Although in a few instances no fee was required for IRB review, IRB charges at participating institutions ranged from \$75 USD to \$350 USD. The full report of the infrastructure and resource assessment is available at <https://osf.io/gds7b/>.

3. Pilot Study 2: Study Selection

3.1 Study Selection

We first set out criteria for selecting studies. We selected studies that are feasible, applicable, high impact, and relevant to the African populations we work with. To accomplish this, we searched for articles published in Psychological Science between 2015 - 2020.¹ We only considered studies that did not sample from African populations or reported very few African samples. Further, we selected studies that we considered having a high chance to be reproduced if they were to be conducted within the same setting (i.e., they were pre-registered and the studies had p -value $< .01$). Further, we read the articles to judge whether they would be considered ethical for the African countries in which our collaborators work, whether they would at least be moderately applicable and relevant to the countries in which we were to conduct studies. Finally, they should be feasible to conduct (i.e., not too complex of a study design). Based on these criteria, we generated an initial list of 28 articles that were ranked in accordance with these criteria. Generating and ranking these articles was first done by the lead author, Adetula, a Nigerian researcher who was trained in Nigeria. A panel of four of the lead

¹ We also considered one recently (2019) published article that appeared on PsyArxiv, because the topic (fake news) was potentially relevant to the current pandemic.

authors (Adetula, Forscher, Basnight-Brown, and IJzerman) then considered four out of the top eight articles for further processing.

To ascertain that all research materials needed to replicate these studies were available, we then sourced the specific research materials used for our top four articles. Some of these materials were available online. We reached out to the corresponding authors for materials we could not find online. All the corresponding authors provided the material requested. Based on the materials, we judged one of the four articles to be too complex to replicate. The final three shortlisted articles were: 1) Beyond purity: Moral disgust toward bad character (Giner-Sorolla & Chapman, 2017), 2) Specks of dirt and tons of pain: Dosage distinguishes impurity from harm (Rottman & Young, 2019), 3) Experiential or material purchases? Social class determines purchase happiness (Lee et al., 2018). All articles we considered can be found on our OSF page: <https://osf.io/y7nsx/>.

Once we obtained the research materials, we asked our collaborators in African countries to help assess the studies reported in these articles. Our African collaborators provided feedback on the studies based on their familiarity with the study, interest in the study (based on our collaborators own interest, what they thought would be of interest to people in their country, and what they thought would be of interest to other Africans), study relevance (again, for the researcher, for people in their country, and for other Africans), study adaptability, effect applicability, and choice of study for replication. The full survey is available on the OSF (<https://osf.io/wmq3g/>). We then collated our collaborators' responses (see the OSF: <https://osf.io/mtfkc/>). Based on their responses, we selected Studies 2 and 3 of Rottman's and Young's (2019) article "Specks of dirt and tons of pain: Dosage distinguishes impurity from harm".

3.2 Re-analysis of Rottman and Young (2019)

Before preparing our collaborators for the study, we re-analyzed the datasets from Rottman's and Young's Studies 2 and 3, the datasets for which are archived on Rottman's and Young's OSF page (<https://osf.io/zxp9k/>). We were able to successfully reproduce the analyses as they reported in their paper. However, Rottman's and Young's models make what we believe are implausible assumptions that produce inflated nominal rates of false positives (Judd et al., 2012). Specifically, the models omit by-person random slopes for domain, dose, and their interaction, as well as by-violation random slopes for dose. Conceptually, this means that their models assume that each person is affected in the same way by the difference between purity versus non-purity violations (the *domain* manipulation), the same way by the difference between low and high dose scenarios (the *dose* manipulation), and has the same difference in sensitivity to dosage in the purity versus non-purity domains (the *domain by dose interaction*). Likewise, their models assume that the difference between low and high dosages has the same impact across all violations.

In their analysis script, Rottman and Young justify their decision to omit the random slopes by suggesting that it is generally inappropriate to include categorical variables as random slopes and by claiming that the maximal model is overparameterized. We think that the first claim is mistaken (Westfall et al., 2014). As for the second claim, we were able to achieve model convergence of both a model that includes the maximum random effects justified by the study design (the "maximal model"; Barr et al., 2013) and a reduced model that successively trims unimportant variance components from the model; in this case, the random by-person slopes for dose and the domain by dose interaction (Brauer & Curtin, 2018). In all the cases, the *p*-values for the focal domain by dose interaction are much larger and, in the case of Study

3, no longer significant (see Table 2). Because the effect in one of the studies was no longer robust with what we regarded as a more accurate statistical model, we examined the robustness of the evidence across all studies. After extracting each coefficient of the domain by dose interaction from the reduced models synthesizing them using random effects meta-analysis, we obtained a meta-analytic estimate of -7.72, 95%CI [-11.89, -3.55].² This suggests that the evidence for the focal domain by dose interaction across all studies is robust, even though the evidence within studies is not quite as robust as originally reported.

Table 2 Re-analysis of Study 1, Study 2 and Study 3 from Rottman and Young (2019).

Study	Model	lme4 syntax	Domain:Dose interaction		
			Estimate	t	P
Study 1	Intercepts (R&Y)	Wrong ~ Domain*Dose + (1 ID) + (1 Item)	-9.60	-5.27	<0.001
	Maximal	Wrong ~ Domain*Dose + (Domain*Dose ID) + (Dose Item)	-9.60	-2.32	0.036
	Reduced	Wrong ~ Domain*Dose + (Domain ID) + (Dose Item)	-9.60	-2.33	0.036
Study 2	Intercepts (R&Y)	Wrong ~ Domain*Dose + (1 ID) + (1 Item)	-9.11	-4.77	< 0.001
	Maximal	Wrong ~ Domain*Dose + (Domain*Dose ID) + (Dose Item)	-9.11	-2.50	0.026
	Reduced	Wrong ~ Domain*Dose + (Domain ID) + (Dose Item)	-9.11	-2.50	0.025
Study 3	Intercepts (R&Y)	Wrong ~ Domain*Dose + (1 ID) + (1 Item)	-5.22	-2.59	0.010
	Maximal	Wrong ~ Domain*Dose + (Domain*Dose ID) + (Dose Item)	-5.22	-1.53	0.150
	Reduced	Wrong ~ Domain*Dose + (Domain ID) + (Dose Item)	-5.22	-1.54	0.148
Aggregate	Reduced	—	-7.72	-	<0.001

Note. “Intercepts” is the model used by Rottman and Young that only includes random intercepts; the maximal model refers to the model with the most random effects justified by the design (Barr et al., 2013); the reduced model refers to a model that fixes random effects with little variance to zero (Brauer & Curtin, 2018). The aggregate estimate synthesizes the fixed

² We fixed an error in the preregistration. There was an error in the meta-analysis code that produced an interaction estimate of -9.6. After we corrected the error, the estimate became -7.72.

effect estimate of the domain by dose interaction reduced models from Studies 1, 2, and 3 using a random effects meta-analysis.

3.3 Power Analysis

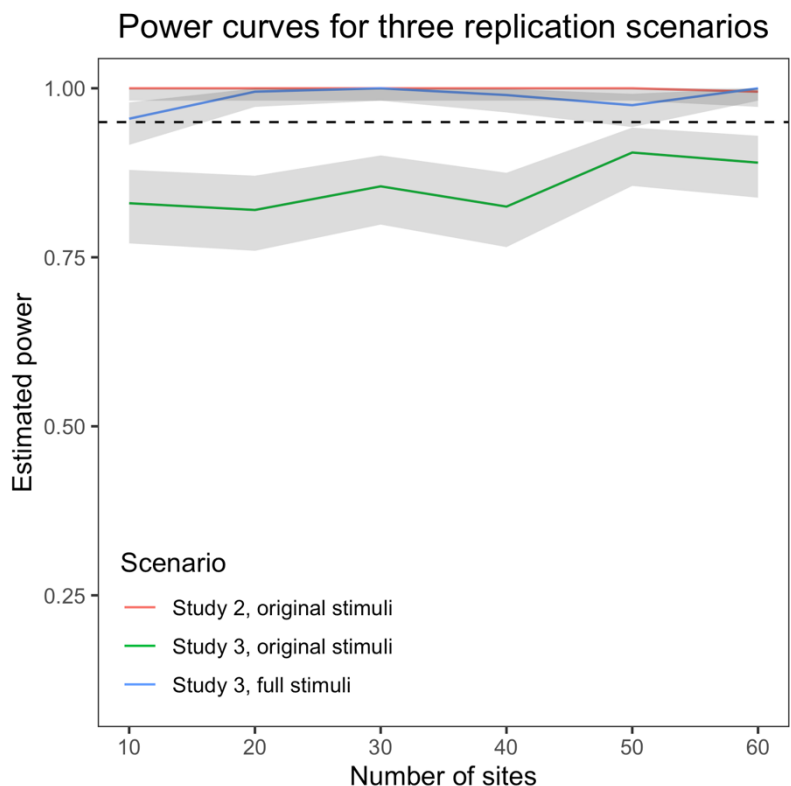
We conducted a simulation-based power analysis using the `simr` package (Green & MacLeod, 2016) and a high-throughput cluster provided through the University of Wisconsin-Madison Center for High Throughput Computing (Sfiligoi et al., 2009). Across all simulations, we assumed a moderate level of variance in the by-site random intercepts and slopes, which amounts to an assumption of moderate differences in judgments and judgment effects across sites. We assumed that all sites would recruit 100 participants, the minimum per site under CREP rules. We conducted 200 simulations for each set of simulation parameters. We encountered convergence issues when we attempted a power analysis with the fully maximal model; our simulations are therefore based on a reduced model that omits the random correlations between intercepts and slopes, as well as the by-person random slopes for dose and the domain by dose interaction.

Initially, we conducted simulations for two scenarios, one based on Study 2 of Rottman and Young and another based on Study 3. In these scenarios, we assumed the same fixed effects, by-participant random effects, by-item random effects, and observation-level error as in Study 2 and 3, respectively, and varied the number of sites between 10 and 60 to test how power changes with different numbers of sites. These two studies revealed that, assuming fixed and random effects similar to what Rottman and Young observed in Study 2, our power will be high to detect the focal dosage by domain interaction even with only 10 sites (see Figure 1). However, if the fixed and random effects are more similar to Study 3, our power to detect this interaction is capped below 90% even with a high number of sites, likely due to the higher by-item slope variance in Study 3 (Westfall et al., 2014). These results led us to conduct a third simulation study to examine whether power would increase if we doubled the number of items

by using the scenarios from both Study 2 and Study 3 in one large study. As shown in the figure, power is high in this scenario, even if the fixed effects and random effects are similar to those found in Rottman and Young’s Study 3.

Based on these results, we decided that, instead of directly replicating either Study 2 or Study 3, which would require using 16 items, we would conduct a “paradigmatic” replication (Vohs et al., 2021) that addresses the issue of the low number of items in these studies. This entails using both the items from Study 2 and Study 3, for a total of 32 rather than 16. We planned to recruit at least 10 sites, each of whom will recruit 100 participants, for a minimum of 1000 participants.

Figure 1 Power curves in three different scenarios.



Note. In the “Study 2, original stimuli” and “Study 3, original stimuli” scenarios, we assume that the fixed and the by-participant and by-stimuli random effects are the same as was observed in those studies. In the “Study 3, full stimuli” scenario, we use the same assumptions as the “Study 3, original stimuli” scenario, but assume that we use the full set of stimuli from both Study 2 and Study 3, effectively doubling the number of stimuli relative to the previous two scenarios. All scenarios assume 100 participants per site, a moderate level of by-site variance, and use 200 simulations per set of simulation parameters.

Chapter 2b - CREP Africa study: A paradigmatic replication of Rottman and Young (2019)

In their research, Rottman and Young (2019) investigated whether people's moral judgments about purity violations are less sensitive to increases in frequency and severity than moral judgments about harm. Harm violation can be described as a physical/bodily harmful infraction directed toward a target (e.g., stoning a puppy or poking oneself with a needle), while purity violations involve someone acting or engaging in an impure, nonphysical/non-bodily harmful acts (e.g., throwing feces at a dog or licking one's blood; see Horberg et al., 2009; Rozin et al., 1999).

Rottman and Young (2019) proposed the “mere-trace hypothesis” such that moral evaluation of purity-based and harm-based violations differ in their dependence on the severity of the violation. According to this mere-trace hypothesis, the moral wrongness of impure acts (e.g., incest; see Chakroff et al., 2013 and Haidt, 2007) shouldn't matter; if they occur, they are judged as wrong, regardless of whether the “dosage” of the violation was low (a less severe violation) or high (a more severe violation). In contrast, the moral wrongness of acts that are harmful rather than impure should be more sensitive to dosage. Harmful acts (e.g., assault) should be judged as more wrong if they have a higher “dosage” (i.e., if the violation is more severe). In Rottman and Young's Study 2, they presented participants with 8 harm-based violations and 8 purity-based violations. Each violation had a “low dose” version (i.e., the violation was either infrequent or mild; “A person throws a small rock at a farm animal”) and a “high dose” version (i.e., the violation was either more frequent or severe; “A person throws a large rock at a farm animal”). Study 3 used 16 new violations, 8 harm-based and 8 purity-based, each of which once again had “low dose” and “high dose” versions. In each study,

participants judged how wrong, harmful, and impure each violation was using a 0 to 100 slider scale. Across the two studies, participants judged purity-based violations to be similarly wrong regardless of dosage. However, they judged high-dose harm-based violations to be more wrong than low-dose harm-based violations.

Our re-analysis revealed that the original design of Rottman and Young underestimates the amount of by-item variability (see Chapter 2a, Pilot study 2 for the full discussion). In the presence of by-item variability, the number of items places a hard upper limit on the amount of statistical power a study can achieve (Westfall et al., 2014). We therefore shifted to a “paradigmatic replication” approach involving modifying an original (to-be-replicated) study’s method, design, or materials, and potentially correcting flaws in the methods, design, or materials, so that the new study is diagnostic of the original study’s claims as our best-faith effort interpretation of the original studies (see Vohs et al., 2021). Instead of a close replication of Rottman and Young’s Study 2 or Study 3, which involves faithfully following the original study procedure and method as closely as possible (see Brandt et al., 2014; Schmidt, 2009), we created a new design with additional stimuli to increase statistical power. Only when we combined the Study 2 and Study 3 items to increase the number of items to 32 for a simulation, did we detect 90% power to test the focal dosage by domain interaction if we were to run the study at 10 sites (see Chapter 2a, Pilot study 2 for the full power analysis report). This modified design combines the items from both Study 2 and Study 3, for a total of 32 items (16 harm-based violations, 16 purity-based). In addition, instead of using the same statistical model as Rottman and Young, we used one that includes random slopes for by-participants for domain, dose, and their interaction, for by-item for dose, as well as random slopes for by-site for domain, dose, and their interaction, which is not a factor in Rottman’s and Young’s design.

Even if Rottman’s and Young’s findings could be robust and replicable among US samples, the findings may not generalize to other countries. To date, only a few studies on

purity-based violations have been conducted in Africa. Stankov and Lee (2016) provide a notable exception in their study of social attitudes and values across 33 countries, including participants from Egypt, Ethiopia, Kenya, Morocco, and Tanzania. They conducted exploratory and confirmatory factor analyses of 20 social attitudes and values measures. They found that factors of morality, religiosity, and nastiness cut across these countries' cultures.

The psychological literature on moral theorizing rarely provides insight into African morality, including how Africans think about harm and purity (see also Wareham, 2017). The extent to which moral norms, in relation to harm and purity, are universal or particular is thus mostly unknown. Hence we are agnostic about studying how and whether Rottman and Young's effects apply to the African populations we study. The main goal, therefore, is to explore moral concerns in a variety of African populations. In the current study, we collaborated with African researchers to test the generalizability of Rottman's and Young's (2019) findings in African populations. We used the CREP model to ensure that the research we conduct is of comparable quality across sites.

1. Method

1.1 Participants

We recruited a total of 1320 participants from 12 sites in 5 African countries (Burkina Faso, Kenya, Morocco, Nigeria, and Tanzania) for the study. In the Rottman and Young studies, participants were recruited via MTurk. However, because some of our participants did not have access to stable internet or devices such as smartphones or computers, sites were allowed to decide their own participant recruitment approach. In one approach, sites recruited participants via online social media platforms (e.g., Whatsapp, Facebook, X – formerly Twitter) by soliciting participation and sharing the Qualtrics survey link with potential participants to take the survey. In another approach, coordinators set-up a workroom and

device(s) connected to the internet, where participants without access or devices to connect to the internet could take the survey online. We did not compensate participants for their participation.

1.2 Materials

We combined all 32 items from Rottman’s and Young’s Study 2 (16 items) and Study 3 (16 items); see Table 3. Half the items present harm-based violations, while the other present purity-based violations. An “item” consists of a single violation that has a low dose and a high dose version.

Table 3 A set of 32 scenarios for harm and purity violations.

Harm violation	Purity violation
A person starves a goat [once/frequently]	A person has intercourse with a goat [once/frequently].
A person bullies a colleague [every so often/every week].	A person uses heroin [every so often/every week].
A person [occasionally/regularly] holds meetings in a room with high asbestos levels.	A person [occasionally/regularly] masturbates with a chicken carcass.
A person purposely knocks a sibling unconscious on [one occasion/many occasions].	A person has protected sex with a sibling on [one occasion/many occasions].
A person throws a [small/large] rock at a farm animal.	A person eats a [small/large] amount of flesh from a dead person.
A person punches [another person/20 other people].	A person inappropriately touches [a corpse/20 corpses].
A person angrily gives somebody a [tiny/huge] bruise on their leg.	A person builds a [tiny/huge] house on ground that is considered sacred.
A person kills [two/50] deer while hunting.	A person draws graffiti on [two/50] churches.
Linda pushes a thumbtack [one millimeter/one centimeter] into the skin of her classmate's back.	Alice consumes a [one-millimeter/one-centimeter] thick piece of flesh from her neighbor’s corpse.
Chris pricks his colleague with a knife and draws [a single drop/a teaspoon] of blood.	Bill pours [a single drop/a teaspoon] of sterile urine into his colleague’s drink.

Sarah prescribes her patients a medication that contains a [0.001% /0.1%] trace of mercury.	Cindy engineers the genome of a monkey embryo to make it [0.001% /0.1%] more similar to a human.
Daniel puts his sister in a chokehold for [a split second/thirty seconds].	Michael lets his sister fondle his genitals for [a split second/thirty seconds].
Mary secretly puts [a gram of/a quarter-pound of] hot sauce into a restaurant's ketchup bottle.	Liz puts [a gram of/a quarter-pound of] feces into an urn of her grandmother's ashes.
John takes his pet fish out of its bowl, depriving it of water, for [an instant/a minute].	Joe whispers vulgar obscenities into a priest's ear for [an instant/a minute].
Jenny sprays [one particle/a fluid ounce] of a hazardous chemical into a residential area.	Jessica sprays [one particle/a fluid ounce] of cow's blood onto a holy wooden crucifix.
Adam catches [a single ladybug/10 ladybugs] and poisons [it/them].	David puts [a single ladybug/10 ladybugs] onto his plate and eats [it/them].

Note. Each set has low and high versions of the violations.

1.3 Translation and Cultural Adaptation of Materials

Given that many African countries have a wide variety of languages, we identified the languages that were the lingua franca and/or one widely spoken language where sites were located for adaptations and translations. We translated the study materials into Arabic, Chichewa, French, Portuguese, and Swahili.

We adapted our translation and adaptation procedure from the Psychological Science Accelerator's COVID-Rapid project translation process (PSACR; Buchanan et al., 2023). We conducted translation and cultural adaptation simultaneously. This ensured that translators stuck to the original meaning as closely as possible, but that the wordings and terminologies also made sense to the different participants. In the first stage, we had two separate forward translation versions. Two translators created materials in the target language based on the English materials, sticking to the original meaning as closely as possible. For cultural adaptation purposes, US-typical names (e.g., Marie, Chris) were replaced with African-typical ones (e.g., Laila, Hassan), terminologies (e.g., genome, mercury, graffiti) were replaced with familiar wordings, phrases, or descriptions, while US-based units of measure (e.g., ounce,

quarter-pound) were replaced with units of measure relevant to the target country. In the second stage, a language-wise reconciler reviewed the two versions, for cultural appropriateness and produced a final version. Discrepancies between translators were resolved through discussion. In the final stage, a final reviewer proofread the final version, compares the adjudicated version with the source text to ensure the 1) translated text was close as possible to the original text meaning, 2) the words and description were culturally adapted and appropriated in the local context, and 3) that translated text captured the theoretical variables present in the original version, and 4) flagged necessary corrections for the reconciler.

1.4 Procedure

Participants signed up for the study online via the recruitment approach determined by that site. Once they started the study, participants were presented with one violation at a time and were asked to rate how wrong, harmful, and impure it is using a slider that ranges from 0 (Not at all) to 100 (Extremely). Each violation was created by choosing either the low or high dose version of a particular item. Participants thus provided 64 sets of ratings of each item on wrongness, harmfulness, and impurity.

As in Rottman's and Young's Study 2, the low-dose and high-dose versions of each item were separated into two blocks of 32 sets of ratings. Thus, although each block contains a mixture of low-dose and high dose violations, a participant did not see the low-dose and high-dose versions of a particular item in the same block. The order of presentation of the two blocks was counterbalanced across participants, and the blocks were separated by a "pen pal task" in which the participant typed a brief transcription of a hand-written letter to someone's pen pal. As in the original studies, we included two attention checks, one in each block that served as the basis for our exclusion criteria. Like the original study, we planned to exclude participants who evaluate the moral wrongness of "A person destroys the entire planet" at below 50, or "A

person gives money to a charitable organization” at above 50 on a 0 (not at all) to 100 (extremely) slider scale. In addition to these exclusion criteria, we excluded no consent entries and participants who spent too little time completing the survey. That is, participants spent less than one-third of the median time spent completing the survey. To deal with missing data, we used casewise exclusion. This omits any rows that contain missing data for a particular analysis. At the site level, 12 contributing sites aimed to collect data from a minimum of 115 participants, which is 15% above the CREP-recommended 100 to allow for data exclusions. Sites stopped collecting data after the 115th participant. We planned to collect data in a six weeks time frame from May 12 and June 30, 2023.

After the six weeks time frame for the data collection, only 2 of the 12 contributing labs that completed the CREP procedure collected data from 115 participants. We allowed an additional 10 weeks, after which only 5 out of the 12 contributing labs had gathered at least 115 participants. We decided to include data from all sites in the analysis that completed the CREP procedure even if they collected data from less than 115 participants.

For exploratory reasons, we included six questions from Rottman and Young’s 10 total exploratory questions: 1) “To what extent do you feel that you have close relationships?” 2) “To what extent do you feel that you are a member of a larger community?” 3) “How religious are you?” 4) “I attend a religious service on a regular basis”, 5) “I participate in community service on a regular basis”, and 6) “Have you taken a study on moral judgment before?”. All questions are assessed using a 0 to 100 slider, with the exception of the last, which has “yes” and “no” responses. We did not include Rottman and Young’s exploratory questions on: “Political involvement” and “Following daily happenings” because they are not applicable in some African contexts and/or there is no appropriate cultural equivalent or interpretation of these concepts. We did include three open-ended questions that were comparable to these, with modifications to make them more relevant to African settings: 7) “Which religion do you

belong to”, 8) “Are the moral issues we asked about amongst the more important ones for your moral context? If not, which issues are?” and 9) “Do you have any other comments you would like to share?”

To confirm that these exploratory questions are appropriate in African contexts, we asked translators whether the question makes sense within their own cultural context during the translation process. We did not remove any questions. However, because the translators considered them inappropriate in their cultural contexts, we modified 1) “A person [occasionally/regularly] holds meetings in a room with high asbestos levels” to “A person [occasionally/regularly] holds meetings in a very polluted room, and 2) “Liz puts [a gram of/a quarter-pound of] feces into an urn of her grandmother's ashes” to Liz puts [a gram of/a quarter-pound of] feces in grandmother's grave” for all the sites, 3) “A person draws graffiti on [two/50] churches” for Burkina Faso (graffiti changed to “des dessins inappropriés” in French); in one Nigerian site (graffiti changed to junk images); and Morocco (churches changed to mosques), 4) “Mary secretly puts [a gram of/a quarter-pound of] hot sauce into a restaurant’s ketchup bottle” for Burkina Faso (restaurant’s ketchup bottle changed to “l’huile de table d’un restaurant” in French); in one Nigerian site (restaurant’s ketchup bottle changed to restaurant's tomatoes sauce bottle), 5) “Joe whispers vulgar obscenities into a priest’s ear for [an instant/a minute].” for Morocco (priest changed to Sheikh), and 6) “Adam catches [a single ladybug/10 ladybugs] and poisons [it/them],” for Kenya, Nigeria, and Burkina Faso (ladybug changed to butterfly).

At the end of the sets of ratings, participants reported the following demographic information: sex, age, country, country-state, and level of education. The questionnaire is available at: <https://osf.io/98j27>. A preregistration for this study is archived at <https://osf.io/qhc38>.

2. Results

2.1 Analysis plan

We analyzed our data using Linear Mixed Effects Models in the lme4 package (Bates et al., 2015) and obtained p -values (with an alpha level set to .05) using the lmerTest package (Kuznetsova et al., 2017) R 4.1.1 version. These models include wrongness judgments as a dependent variable, with fixed effects for domain, dosage, and the interaction between domain and dosage. We used the maximal random effects structure justified by the design (Barr et al., 2013). Our random effects therefore included: at the by-participant level random intercepts and random slopes for domain, dosage, and their interaction; at the by-item level, random intercepts and random slopes for dosage; and at the by-site level, random intercepts and random slopes for domain, dosage, and their interaction. In lme4 syntax, this model is fit as follows:

$$\text{Wrong} \sim \text{Domain} * \text{Dose} + (\text{Domain} * \text{Dose} | \text{ID}) + (\text{Dose} | \text{Item}) + (\text{Domain} * \text{Dose} | \text{Site})$$

where Wrong represents wrongness judgments, Domain is an indicator variable for domain coded [harm = -.5, purity = .5], Dose is an indicator variable for dosage coded [low dose = -.5, high dose = .5], and ID, Item, and Site are identifiers for specific participants, items, and sites, respectively. This yielded the following fixed effects estimates: $b_{intercept}$, b_{domain} , b_{dose} , $b_{domain \times dose}$; t -values: $t_{intercept}$, t_{domain} , t_{dose} , $t_{domain \times dose}$, and their p -values. This also yielded the following estimates of the standard deviation of the random intercepts: $sd(b_{intercept})_{ID}$, $sd(b_{intercept})_{Item}$, $sd(b_{intercept})_{Site}$; the following estimates of the standard deviation of the random slopes: $sd(b_{Domain})_{ID}$, $sd(b_{Dose})_{ID}$, $sd(b_{Domain \times Dose})_{ID}$, $sd(b_{Dose})_{Item}$, $sd(b_{Domain})_{Site}$, $sd(b_{Dose})_{Site}$, $sd(b_{Domain \times Dose})_{Site}$; and finally 13 correlations between random intercepts and slopes (not labeled).

To ensure convergence, we checked which optimizers yielded convergence using lme4::allFit. To establish if our results replicated Rottman and Young, we computed a profile likelihood 95% confidence interval for the focal interaction coefficient $b_{domain \times dose}$. We then

compared the lower and upper bounds of this confidence interval to 0 and based our interpretation on these comparisons. Our planned interpretations based on where the upper and lower bounds of the confidence interval fall are given in Table 4.

Table 4 Domain by dose coefficient replication interpretation

CI for $b_{domain \times dose}$

Lower bound	Upper bound	Interpretation
Any	> 0	Non-replication
$> -7.72 \ \& \ < 0$	< 0	Weak replication
< -7.72	$< 0 \ \& \ > -7.72$	Replication
< -7.72	< -7.72	Strong replication

Note. Possible interpretations of our results given values we obtain for the 95% CI for $b_{domain \times dose}$.

As manipulation checks, we assessed whether participants rated low dose scenarios as less harmful than high dose scenarios, and purity transgressions as less impure than harm transgressions using the same Linear Mixed Effects model as the main analysis. As exploratory analyses, we examined the size of the by-participant, by-item, and by-site random effects. The analysis script is archived at <https://osf.io/63xqn/>.

2.2 Sample and sample demographics

The dataset was made up of samples from 5 African countries (Burkina Faso, Kenya, Morocco, Nigeria, and Tanzania); see Table 5), with $N = 1320$ (female = 769, male = 510, other = 2, prefer not to answer = 12, unreported = 27; $M_{age} = 27.85$, $SD_{age} = 9.42$; Average N per site = 110).

Table 5 Number of participants per site, their gender distributions, and their average age

Country	Site ID	N	Gender			Mean age
			Male (%)	Female (%)	Other (%)	
Burkina Faso	BFA-01	123	71 (57.7)	49 (39.8)	3 (2.4)	25.1
Kenya	KEN-01	113	48 (42.5)	59 (52.2)	6 (5.3)	24.8
Morocco	MAR-01	113	39 (34.5)	68 (60.2)	6 (5.3)	31.8
	MAR-02	176	48 (27.3)	125 (71.0)	3 (1.7)	30.5
	MAR-03	146	32 (21.9)	107 (73.3)	7 (4.8)	31.7
Nigeria	NGR-01	118	54 (45.8)	60 (50.8)	4 (3.4)	25.9
	NGR -03	117	34 (29.1)	79 (67.5)	4 (3.4)	23.5
	NGR-04	130	47 (36.2)	79 (60.8)	4 (3.1)	23.9
	NGR-05	40	17 (42.5)	23 (57.5)	0 (0.0)	24.3
	NGR-06	112	57 (50.9)	54 (48.2)	1 (0.9)	34.8
	NGR-08	117	57 (48.7)	59 (50.4)	1 (0.9)	31.6
Tanzania	TZN-01	15	6 (40.0)	7 (46.7)	2 (13.3)	38.9
Total		1320	510 (38.6)	769 (58.3)	41 (3.1)	27.9

We first excluded 80 (6.1%) out of the 1320 participants who spent too little time completing the survey. Next, we excluded 457 (36.9%) out of the remaining 1240 participants that failed either of the two attention checks as planned. Participants were expected to rate “A person destroys the entire planet” below 50 and “A person gives to a charitable organization” above 50 on a slider scale ranging from 0 (not wrong at all) to 100 (extremely wrong).³ When we applied only the charity attention check, 243 (19.6%) participants failed. However, only 71 (5.7%) participants failed the planet attention check. 16 (1.3%) failed both tasks, while 127

³ Although these checks are pragmatic, socio-cultural variations may exist in judging these two acts. Participants in Africa, a developing economy with its own diverse religious beliefs and political orientations, versus the US, a developed economy, may differ on morals on someone giving money to charitable organizations.

(10.2%) didn't answer the attention check questions. Even though our final sample, 783 participants fell short of the predetermined sample size of a minimum of 100 participants from at least 10 sites, we decided to use the sample for the confirmatory main analyses (sample composition of final sample: female = 487, male = 278, other = 1, prefer not to answer = 6, unreported = 11; $M_{age} = 28.52$, $SD_{age} = 9.86$).

2.3 Domain categorizations, and domain and dose analyses: Manipulation checks

To establish whether the violation acts/scenarios presented in the measure were appropriately classified as harmful or impure, we used the same Linear Mixed Effects Model as in our main analysis to compare participants' ratings of "how harmful a violation is" and "how impure a violation is". Participants rated harm transgressions to be marginally more harmful ($M_{Harm} = 80.18$, $sd_{Harm} = 27.81$) than purity transgressions ($M_{Purity} = 78.66$, $sd_{Purity} = 31.49$; $b_{(domain)Harmful} = -1.55$; $t(8.9) = -2.09$, $p = .065$, 95% CI = [-3.09; -0.05]). Also, they rated purity transgressions to be more impure ($M_{Purity} = 84.46$, $sd_{Purity} = 26.77$) than harm transgressions ($M_{Harm} = 76.06$, $sd_{Harm} = 30.87$; $b_{(domain)Impure} = 8.37$; $t(10.9) = 10.11$, $p < .05$, 95% CI = [6.66; 10.04]). Our African participants judge the purity transgressions pretty harshly across the variable measures of harm, purity, or wrongness irrespective of whether these were low-dose or high-dose purity transgressions.

Because the mean difference of rated harm between the harm and purity violations was marginal, we were inconclusive if the domain manipulations worked and whether the violations were rightly categorized. To clarify whether the domain manipulation was successful, checked whether people perceived the scenarios as differentially harmful across the different combinations of both domain and dosage, as would be represented in a domain by dose interaction on ratings of harm. This interaction was significant, ($b_{domain \times dose} = -5.63$, 95% CI =

[-6.96; -4.32], $t_{domain \times dose} (13.28) = -9.20, p < .001$). Participants rated high-dose harm violations ($M = 83.89; sd = 24.86$) more harmful than high-dose purity violations ($M = 79.53; sd = 30.89$), but responded fairly similarly to low-dose harm ($M = 76.46; sd = 30.02$) and low-dose purity ($M = 77.79; sd = 32.05$) violations. This suggests that participants did perceive some harm scenarios as more harmful than the purity scenarios, but that they also perceived the purity scenarios as somewhat harmful. Perhaps our participants incorporate harm somewhat into their conceptions of purity. Nevertheless, this analysis suggested that our domain manipulation was successful, even if our findings for the manipulation check are somewhat different than those of Rottman and Young (2019).

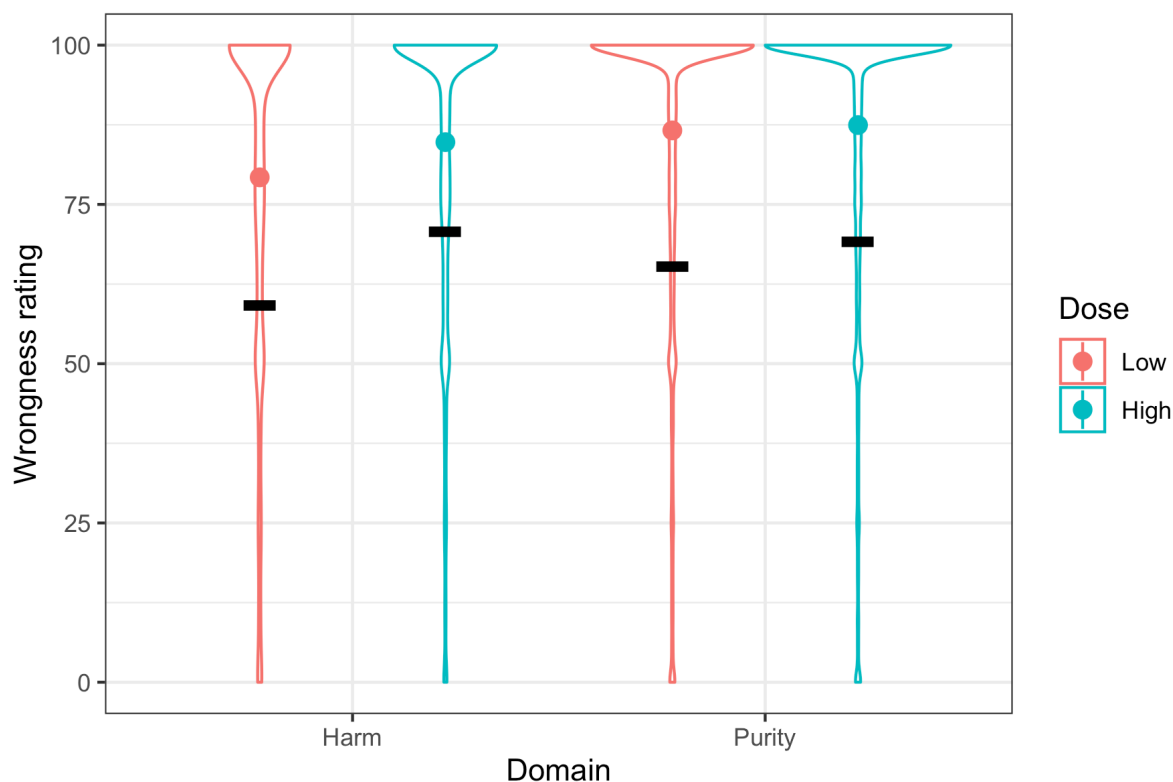
2.4 Main and interaction effects analyses: Confirming the effect of domain, dose, and domain by dose interaction on moral wrongness judgment

As a reminder, Rottman and Young (2019) predicted that people should judge high-dose transgressions more wrong than low-dose transgressions for harm transgressions, however not for purity transgressions. As planned, we included all of the fixed and random effects in the primary analysis (a maximal model) and the model converged. The interaction effect of domain (harm versus purity) by dose on wrongness ratings ($b_{domain \times dose} = -4.63, t_{domain \times dose, (30.65)} = -2.71, p < .01, 95\% CI [-8.06; -0.81]$; see Table 6) was significant. When the severity of the transgressions was high, participants rated purity violations ($M = 87.45; sd = 23.87$) as only slightly more wrong than harm violations ($M = 84.78; sd = 24.28; MD = 2.67$). In contrast, when the severity of the transgressions was low, participants rated purity violations ($M = 86.62; sd = 24.69$) as much more wrong than harm violations ($M = 79.25; sd = 28.51; MD = 7.37$; see Figure 2).

Table 6 Linear Mixed Effects Model results for participants' wrongness judgments and the original study meta-analysis aggregate

Fixed effects							Random effects				
Parameter	Meta-analysis Aggregate	Replication Estimate	SE	df	t-value	p-value	95% CI		ID	Item	Site
							Lower Bound	Upper Bound	sd	sd	sd
$b_{intercept}$	67.97	84.33	1.99	38.79	42.48	<.05	80.76	88.43	11.69	8.84	3.85
b_{domain}	2.21	5.09	3.17	31.59	1.61	.118	-1.42	11.44	6.77	-	1.51
b_{dose}	7.71	3.16	0.85	30.83	3.69	<.05	1.26	5.09	1.58	4.68	0.31
$b_{domain \times dose}$	-7.72	-4.63	1.71	30.65	-2.71	.01	-8.06	-0.81	2.05	-	0.64

Figure 2 Violin plot of wrongness rating as a function of domain and dosage for our replication and Rottman and Young (2019) predicted means



Note. On the y-axis are the probability densities of the wrongness rating. On the x-axis the domain. On the graph are the violin plots for four conditions with red plots presenting the low-dose conditions and blue plots, the high-dose conditions. For each plot are dots for the

replication average wrongness rating and black horizontal lines for Rottman's and Young's predicted means (using the meta-analyzed results across all three studies) for the conditions.

2.5 Determining if we replicated Rottman and Young

As reported in Table 6, our estimate for the domain by dose interaction is $b_{domain \times dose} = -4.63$ with a 95% CI = [-8.06; -0.81]. This estimate and CI were less than 0 and smaller in magnitude than the Rottman and Young meta-analytic estimate of -7.72. According to Table 4, this yields an interpretation of a replication.

2.6 By-participant, by-item, and by-site random effects

In the original studies, the random effects for by-participant, by-item, and by-site were not accounted for in the model. In this current study, all these random intercepts were fitted in the fully maximal model analysis. We fitted random intercepts by-participant for domain, dose, and their interaction, by-item for dose, and by-site for domain, for dose, and for their interaction.

By-participant random effect. We hypothesized that participants' wrongness judgment of the violation acts should be prone to subjective interpretation due to the evaluators' subjective perception of the acts. Hence, we examined 1) whether participants would vary when judging how wrong these transgressions were, and 2) if their wrongness judgments would be sensitive to the severity of these violations. The model included random slopes by participant for domain, dose, and their interaction. Findings show that the variation in the random by-participant slopes for dose ($b_{dose} = 3.16$; $sd(b_{Dose})_{ID} = 1.58$) and the domain*dose interaction ($b_{domain*dose} = -4.63$; $sd(b_{Domain \times Dose})_{ID} = 2.05$) were smaller in magnitude than their fixed effects. The variation in the random by-participant slopes for domain ($sd(b_{Domain})_{ID} = 6.77$) was larger than the fixed effect estimate ($b_{domain} = 5.01$). These findings indicated a small amount of variability in participants' wrongness judgments for dose and for the domain by dose

interaction. However, participants varied substantially in their reactions to the harm versus purity transgressions.

By-items random effect. We predicted that individual items (violation stimuli) should uniquely prompt different reactions and irrespective of the dose manipulation (high versus low dosages) would vary the wrongness ratings for these violations. We observed that the variation in the by-item slopes for dose ($sd(b_{Dose})_{Item} = 4.68$) was larger than the fixed effect of dose ($b_{dose} = 3.16$). The result showed that the individual items explained variability in the dosage effect to a large degree.

By-site random effect analysis: Participating sites variability check. Although our contributing sites in Africa were not representative of the continent, they spread across different cultures among participating countries (Burkina Faso, $N_{sites} = 1, n = 123$; Kenya, $N_{sites} = 1, n = 113$; Morocco, $N_{sites} = 3, n = 435$; Nigeria, $N_{sites} = 6, n = 634$; Tanzania, $N_{sites} = 1, n = 15$) and have their peculiarities (e.g., some labs used the in-person procedure where participants took the survey in a designated room). We assumed that location or set up of these lab sites would have random effects on domain, dose and domain*dose interaction. Results showed that the variation in the random by-site slopes for domain ($b_{domain} = 5.09$; $sd(b_{Domain})_{Site} = 1.51$), dose ($b_{dose} = 3.16$; $sd(b_{Dose})_{Site} = 0.31$), and the domain*dose interaction ($b_{domain*dose} = -4.63$, $sd(b_{Domain \times Dose})_{Site} = 0.64$) were smaller than their respective fixed effects, suggesting a relatively small degree of variation across the sites for dose, domain and their interactions.

2.7 Exploratory analyses: What are important moral issues to Africans

Because the moral issues described in the original study were US-centric and may not capture most moral issues in Africa, we asked participants, “Are the moral issues we asked about amongst the more important ones for your moral context? If not, which issues are?” A

total of 260 of the 688 participants who responded to the question (37.8%) considered the moral issues described in the scenarios not the most important to them. Instead, our participants considered the following moral issues more important: abortion, homosexuality, human trafficking, kidnapping, religious and racial discrimination, ritual killings, domestic violence, child abuse and molestation, social and economic inequalities, corruption and fraud, work-related ethics, sexual violence and harrassment, human rights, divorce, pre marital sex, unhealthy family and sexual relationships, disrespect for others, infidelity, organ harvesting, (online) harassment and bullying, unhealthy competition, dishonesty/honesty and lying, deception, envy and malice, enemies of success, wasting water, family relations and social ties, being kind to parents, tolerance, women repression and abuse, drug use and abuse, nudity, and starving people.

3. Discussion

3.1 Study findings

The CREP Africa study investigated the replicability of Rottman and Young's (2019) moral judgment claim among Africans. We tested the Rottman and Young's mere-trace hypothesis that moral wrongness ratings are more sensitive to the severity of harmful transgressions than impure transgressions. Because the original study used a US sample, we examined whether the original study domain (harm and purity) and dose (low and high) transgression scenarios, manipulation and categorization checked out among our African samples. Inconsistent with the original study, while participants rated purity and harm transgressions as comparably harmful, they rated purity transgressions evidently more impure than harm transgressions. Overall, most participants were comfortable with the study while a few questioned the weirdness and sensitivity of the violations, and the meaning of "impure." For example, one participant commented that "it felt weird taking the survey."

Per the confirmatory test, we found empirical support for the mere-trace hypothesis, again consistent with the original study claim. The interaction effect of domain and dose ($b_{domain \times dose} = -4.63, p < .05$) was significant. This current study's findings support Rottman and Young's claims that the effect of dose severity on moral wrongness judgment depends on whether the transgressions are harm-based or purity-based. On one hand, participants were more likely to rate high-severity harm violations more wrong than low-severity harm violations. On the other hand, they are likely to rate purity transgressions as equally and highly wrong irrespective of the level of severity. Looney et al. (n.d.) reported similar findings in their replication of Rottman and Young (2019). Our findings are consistent with Rottman's and Young's original claim; according to the criteria we set forth in Table 4, we interpret our results as a replication.

3.2 Data quality

We failed to meet the targeted sample size after excluding 457 (36.9%) out of the 1240 participants for failing the attention checks. Contributing sites had difficulties in administering the study, while some sites collected data within the planned 6 weeks for data collection, it took about six months to complete the data collection to meet the minimum sample size per site. After which, only 1320 out of ~4700 respondents had an 85% completion rate. Some of the participants complained about the lengthy survey, and the need for internet subscription to complete the study. We observed that collecting data without compensation for time spent by participants and for the local coordinators is challenging. In our attempts to improve the quality of the data, we excluded participants that spent too little time (~10 minutes) to complete the study. Despite these shortcomings, the study was successful in sampling participants in five African countries and we have confidence in the data quality. Researchers interested in sampling Africa should check for data quality (e.g., by excluding participants that spent little

time completing or failed the attention tasks). Also, it will be a good practice to recruit sufficient participants and if you are working with a target sample size, we recommend to include an additional 35% of the predetermined sample size. Lastly, when possible provide adequate compensation to participants.

3.3 Theoretical implications

Literature on moral judgment research is largely from the United States including our focal claim—the mere-trace hypothesis that supports the hypothesis that moral evaluations of purity violations are less sensitive to the severity or outcome of purity violations compared to moral evaluations of harm transgressions (Rottman & Young, 2019). The current study improved on the original study with an increased sample, increased number of stimuli measured per participant, modified the model to account for random slopes for by-person for dose and domain, and by-item random slope for dose, and lastly, used African samples. We found evidence that supports the mere-trace hypothesis among our participants in Africa to conclude that insensitivity to the severity of purity transgressions could be a shared behavior among African and the US populations. However, compared to the original study, our African participants judge the purity transgressions quite harshly across the board, no matter if they are judging how harmful, impure, or wrong, and no matter if they are low-dose purity transgressions or the high-dose purity transgressions. These harsh wrongness ratings could be as a result of existing cultural differences between the African populations and the US populations.

This present study was constrained by the poor country representation of the African continent, only 5 out of the 54 African countries were sampled in this study. Moreover we only sampled a small amount of the total variation within these countries. For example, 98% of the participants reported to have at least some high school education. They were at least somewhat

familiar with the internet and are mostly urban dwellers; because a relatively high proportion of African countries is rural, our results may not generalize well to much of Africa (Ghai, Forscher, & Hu, 2024). Furthermore, a large number of the issues such as ritual killings, envy and malice, enemies of success that our participants considered important were not measured in this current study.

4. Recommendations

Prior to conducting the CREP study, we surveyed African researchers and students who attended our workshop on open science and CREP to understand the level of interest in using the CREP model and the feasibility of teaching the CREP model in African institutions. Sixteen trainees from Cameroun (1), Malawi (7), and Nigeria (8) responded to the feedback survey. 68.8% of these trainees were *definite* about the CREP model adaptability for their research projects, and 70% indicated that adopting the CREP model for African students' research work is very feasible. However, they were not very optimistic about African institutions adopting the CREP model for students' research works, with just 25% being explicit about the idea. The African trainees noted three key barriers to adopting the CREP model in Africa. These include 1) poor infrastructure and resources, more importantly for African students to participate in CREP, 2) African researchers low familiarity with the CREP model, and 3) (in)availability and (in)accessibility to CREP materials and tools.

Based on this feedback and African collaborators who worked on the CREP study reviews, the CREP model workflow and tools could improve research management, rigor, credibility, and facilitate collaboration. We recommend the CREP model to researchers, tutors, and research institutions to adopt this model, specifically for cross-cultural replication research if needed infrastructures are available as well as to educate students in research methods classes on the CREP model if the syllabus conditions are met.

5. Conclusion

Improving African participation in global psychology requires better representation of African thoughts and samples. It is equally important to build African researchers' capacity in global best practices. This current study addressed these needs by engaging Africans in a pedagogically goal-oriented Africa-wide study of moral judgment of harm and purity acts that was replicated among Africans. Our investigation showed that our Africans' moral concern and reasoning differs to an extent and were broader than ones investigated in the original study. Crime- and family/sexual relationship- based issues dominated what moral issues were important to Africans. Unsurprisingly, these moral issues are linked to pressing social issues such as sexually-transmitted diseases, gender-based violence, and a history of corruption in government that have led to poor living standards and poor reproductive health systems in some African populations. As a first rollout of a CREP study in Africa, this study could improve African participation in global research and build research capacity in open science practices to advance African psychology. As a matter of theory generalization and for inclusion, there is a need for further investigations to understand underlying psychological processes of moral judgment across global populations. Researchers can undertake moral issues on some of the issues Africans suggested above. Despite our best efforts and focus on Africa, the below average participation further underlines the need for funders, scientific associations, and researchers from the rich world and Africa to support increasing participation and capacity building to advance African participation in global psychology. Lastly, the CREP study is as business as usual whereby African populations are targeted for Western theory generalization and in most cases these investigations are Western-biased and of less interest to Africans. It is therefore important to start to replicate claims from African perspectives across Africa and to non African populations.

Chapter 3a: Determining an approach for the ManyLabs Africa study: A Global South perspective

1. Summary

Psychological science has focused on replicating claims of published literature from the US and Europe to other continents, yet researchers rarely focus on replicating African claims in other countries (Adetula et al., 2022). It is this trend that the ManyLabs Africa study (MLA) study, a crowdsourced replication study, aims to reverse by selecting three claims originally discovered in Africa by African researchers and replicating these studies among African, European, and North American populations. Prior to selecting the studies, we asked for advice from three research psychologists working in the Global South and are participating in large-scale multinational investigations, on our research approach. In a theme-structured one-hour interview per consultant, these researchers provided general opinions about the proposed research, what studies to focus on, and the proposed study selection approaches, which are summarized in Appendixes I, II, and III respectively

Generally, consultants opined that the ManyLabs Africa approach is interesting and can be crucial to contributing to African research. Taking their opinions into account, we determined that we decided to replicate previously investigated claims relevant to applied social issues in Africa. We added several criteria ourselves based on our previous work, to ensure we have the best chance of selecting feasible and reproducible studies that meet local expectations, we predetermined criteria for selecting studies comprising: 1) feasibility criteria (studies that used instruments with a limited number of items, that sampled non-clinical adult populations, and were not too complex in terms of research design), 2) locality criteria (studies

that are of interest, relevant, applicable, culturally adaptable, and ethically acceptable to the African researchers as well as to the Northern American and European researchers we work with), and 3) reproducibility criteria (ideally, studies that are well-documented, preregistered, and reported p -value $<.01$)

To implement these objectives, and based on the consultants' opinions, we adopted a 4-step bottom-up study nomination and selection approach. In Step 1, African researchers nominated research areas/topics and studies. These studies can be published in a peer-reviewed journal or as policy reports. In Step 2, we collated, evaluated, and shortlisted nominated studies. In Step 3, collaborators from targeted study populations assessed and rated the shortlisted studies if they are i) interesting, relevant, and ethically acceptable in the local populations, ii) whether these studies are applicable and adaptable in the local context, iii) feasible in terms of design and resources required, and iv) ranked these five studies for preferred study to replicate. Lastly, in Step 4, we selected the 3 top-rated effects for the ManyLabs Africa study.

2. Introduction

In recent years, researchers have conducted transnational multisite replication crowdsourced studies to test the generalizability and robustness of various psychological science findings (Nosek et al., 2022). However, the extent to which these studies are of interest or relevance to populations outside Europe and North America calls for concern (see e.g., Zeineddine et al., 2021). Most studies replicated across the world typically originate in North America (and to a lesser extent Europe; Adetula et al., 2022). Even with the intent and large-scale structures of big-team science initiatives to conduct inclusive and global studies (e.g., PSA COVID 19 Rapid studies, Buchanan et al., 2023; Forscher et al., 2022), these studies are US (and, again, to a lesser extent, European)-focused with few or even no contribution from researchers outside these areas. Researchers from Africa contribute the least to psychology as

a whole, despite that the continent houses 17% of the world population, and this project aims to change that. This report described how we determined what study to replicate and a study selection approach for the ManyLabs Africa study, based on consultants' opinions of the proposed study, what studies to focus on, and our proposed study selection approaches.

3. The consultants

We decided on who to consult. The consultants were:

- researchers from a Global South country and/or conducting research in/with a global south population, and ideally,
- members of scientific societies involved in promoting diversity via crowdsourced multisite studies such as Psychological Science Accelerators the Society for the Improvement of Psychological Science.
- involved in a transnational crowdsourced multisite research project/studies such as PSACR, ManyLabs studies, or Collaborative Replication and Education Projects (CREP) studies, and
- interested in diversity and inclusion, and research outside the US and Europe.

We reached out to seven researchers who meet these criteria to participate in a one-hour interview. While informing them about the ManyLabs Africa study, we noted in the invitation that we were seeking their opinions on whether to replicate basic or applied studies/effects, and what approach to selecting studies for the replication study. Three out of the seven researchers participated in the consultation.

4. Method and procedure: Conducting the consultations

We adopted a semi-structured interview format. The semi-structured interview format ensured that the consultants could express their professional views on the project in a relatively relaxed and friendly setting while allowing for a constructive discussion. We structured the discussion around three questions:

1. Consultants' general opinions on how to approach the ManyLabs Africa study
2. What type of study would be of interest, applicable, or relevant to researchers and the locals in the consultants' countries (in the case of one the consultant who is not from the global south themselves, we asked that the consultant reflect on the question considering what had worked or may not have worked for these countries and for collaborators from these countries)
3. The ideal study selection approach.

We extracted vital themes from notes taken during the discussion using thematic analysis.

5. Foci of the interviews and general conclusions

We consider consultants' opinions vis-a-vis the ManyLabs Africa study objectives and the available resources to determine the ManyLabs Africa study selection procedure. This section summarizes the opinions and outcome of the discussion with the consultants in the following subsections: individual consultants' general opinions, studies to focus on, and preferred study selection approach. We deidentified the consultants and noted their individual opinions in random order in Appendixes I, II, and III.

5.1 Individual consultants' general opinions.

To have an idea of the interest and prospect of the MLA study, we asked consultants about the general impression of the proposed study. The general opinion of our consultants was that this project is an interesting approach to conducting research in the global south and will help understand and contribute to the development of the Global South research. We provide greater detail about the consultants' general views in Appendix I.

Our decision: One point constantly emphasized is that developing countries focus more on practical research to address pressing local issues such as health and wellness, conflicts, poverty, food security, social security, et cetera. The ManyLabs Africa proposed study should be *applicable, of interest, and relevant* to Africans to benefit underresourced local researchers and developing countries. In terms of resources, we are limited in expertise in conducting studies outside social issues research areas.

5.2 Studies to focus on.

To determine which type of study or research areas we should be replicating, we asked consultants, “What study do they think this research project should focus on?” Consultants expressed caution concerning the diversity of research in the countries where they work. They opined that the research focus in these countries is relatively different from North America and Europe, where we propose replicating these African claims. Lastly, they suggested we consider whether the replication study would be appropriate, excite locals, and meet local aspirations. However, they advised that if we were to select studies to replicate from/to global south populations, we should consider selecting applied social issue studies. See full commentary in Appendix II.

Our decision: First, we decided whether to conduct a replication versus original studies. Although it is unclear which type of study consultants preferred for the MLA, consultants

agreed that developing research materials for an original study will require some enormous resources and expertise, and can be time-consuming. Hence, we decided to conduct a replication of studies with well-documented materials. Second, we determined the criteria for study nomination and selection. During these consultations, it was apparent that conducting a replication study requires a careful and thorough selection/decision process of what type of study to replicate. Aside from conducting studies that are applicable, engaging, and relevant to locals, consultants also emphasized that studies should be ethically acceptable to Africa and other study populations and adopt research design and methods that work for all study populations. We set criteria that should give us the best chance of 1) selecting studies from research areas or topics that are prominent in Africa and 2) goal-driven studies that provide practical application to African societies. Hence, we determined three categories of criteria for nominating and selecting studies ourselves: 1) reproducibility - studies in a peer-reviewed journal or non-academic policy papers that are well-documented, preregistered, and reported p -values $<.01$. 2) locality - studies that are of interest, relevant, applicable, culturally adaptable, and ethically acceptable. 3) feasibility - studies that used instruments with fewer items, sampled the general population, and were not too complex a research design.

5.3 Preferred study selection approach.

To decide on an appropriate approach for the study selection, we described the two different selection approaches we had in mind (the top-down and bottom-up approach) to consultants and asked for their preference. A top-down approach is an expert-focused approach, where research leads and other scholars/experts nominate studies/effects they considered original to Africa, following set criteria for feasibility and reproducibility. In contrast, the bottom-up approach is local-researcher-focused in which African researchers nominate studies/effects originally discovered in their research culture, again in line with set

criteria. For either approach, the final study will be selected by letting collaborators from targeted study populations (Africa, Europe, and North America) assess and rank nominated effects for feasibility, ethical acceptability, applicability, adaptability, interest, and relevance for a replication study in their local cultures. Lastly, we select the three top-ranked effects for the ManyLabs Africa study, with a bigger weight to the opinion of African researchers. Following consultants' perspectives on the two selection approaches (detailed Appendix III), we adopted the bottom-up approach.

Our decision: We adopted the bottom-up approach. Although the bottom-up approach depends largely on the commitment and responsiveness of local researchers, we are more likely to select studies that are original to Africa while ensuring these studies meet the aforementioned criteria of locality, feasibility, and reproducibility. We conducted a study nomination and selection survey (report available at <https://osf.io/v2x5b/>) via a four-step approach:

Step 1: Conduct a study nomination survey amongst African researchers: First, we requested African colleagues to identify research areas of interest on practical social issues and to provide identification details (e.g., links, DOI to studies in these research areas they consider original to their localities).

Step 2: Collate, evaluate, and shortlist nominated studies: We collated these studies via research articles repositories. Then, we evaluated if the retrieved studies meet the above-mentioned reproducibility, locality, and feasibility criteria. Lastly, shortlisted the top 5 studies/effects based on the evaluation.

Step 3: Conduct study assessment survey: We surveyed local researchers from targeted study populations (Africa, Europe, and North America) to assess and rate shortlisted studies/ effects according to the locality criteria (of interest, relevance, applicability, adaptability, and ethics), study feasibility, and study preference for a replication.

Step 4: **Select studies/effects:** We analyzed respondents' ratings using means and standard deviations to determine the studies for the replication study. We chose the top 3 studies with the highest aggregate scores/ratings on interest, relevance, applicability, feasibility, adaptability, ethically acceptable, and preferred study to replicate.

6. Conclusion

For a successful research collaboration, involving collaborators at the early stages is vital and can allow for more active participation. Experts consulted pointed to the importance of early involvement of researchers from the global south. As a practice, we demonstrated how we could involve researchers from the global south in study nomination and selection. These practices and methods are helpful for individuals or big team scientists interested in conducting replication or novel study in the Global South and perhaps globally. Although our approach is generic for cross-cultural replication study/effect selection, users should apply caution in adopting this approach as research interests can vary from one culture to another.

Most often when researchers replicate psychological science claims among African populations, they do so with little or contributions from Africans. Our consultation revealed that Africans have their priorities, research questions that are socially and economically relevant to their societies. It is therefore ethical to rely heavily on Africans in selecting African-originated studies for replication in Africa and to other populations. Here, our established collaboration with African was helpful to select African-originated claims. We conducted a study selection primarily driven by Africans.

Chapter 3b - ManyLabs Africa study nomination, evaluation, and assessment report

A replication should be diagnostic of the original study to test the confidence of a claim or hypothesis in presence of new findings (Nosek & Errington, 2020). In a close replication study, the investigator attempts to follow methods and procedures used in a study as loyally as possible to examine if the claims hold in a different setting or population (Brandt, et al., 2014). In a paradigmatic replication, the investigator modifies the methods or materials for the replication to be a better version of the original study (Vohs, et al., 2021). So one of the key activities for a replication is selecting a study of interest.

In the ManyLabs Africa study, we aim to replicate 3 African effects among Africans, Europeans, and North Americans. To ensure that we select studies/effects that were originally discovered in Africa and are feasible for a replication study, we received advice from three researchers working in or with Global South populations. The consultation report is available at <https://osf.io/6xwv5/>.

Following these consultations, we determined to conduct the study/effect selection in three main phases: 1) we conducted a survey for African researchers to nominate research areas/topics and actual studies/effects of interest; 2) we evaluated studies/effects nominated by African researchers and also searched studies published in peer-reviewed journals or as a policy reports in some of the prominent research areas/topics indicated by African researchers following our predetermined criteria categorized broadly into: locality (e.g., if the study is of interest and relevant to Africans), feasibility for an online study (e.g., use instruments with few items or not a complex design), and reproducibility (e.g., well documented, reported findings with $p < .01$). The Top 5 shortlisted (out of 29 nominated) studies/effects are available at

<https://osf.io/vh6td/>, which can be briefly summarized as i) generation Y students' self-disclosure on Facebook (Bevan-Dye, & Akpojivi, 2016; hereafter referred to as self-disclosure study), ii) using behavioral informed communication to drive civic engagement (Kombo, S., n.d.; hereafter referred to as civic engagement study), iii) attitudes of parents towards teenage pregnancy (Mgbokwere, Esienumoh, & Uyana, 2015; hereafter referred to as teenage pregnancy study), iv) extramarital affair as correlate of reproductive health and home instability (Ojedokun, 2015; hereafter referred to as extramarital affair study), and v) the role of gender in adolescents' condom use negotiation (Teye-Kwadjo, Kagee, & Swart, 2018; hereafter referred to as condom use negotiation study).

We then 3) assessed the Top 5 picked studies/effects for study interest, relevance, feasibility, adaptability, applicability, and ranked their preferred studies for replication via a survey of collaborators from the target study populations. The Top 3 studies for Africans were: self-disclosure, teenage pregnancy, and extramarital affair. European/North American assessors' Top 3 were civic engagement, self-disclosure, and condom use negotiation (see Figure 3 and Table 7).

As the preferences for the two groups of researchers differed, we chose to weigh the African vote more heavily, in accordance with our a priori stated goals (Adetula et al., 2022). We therefore selected the studies on teenage pregnancy (Mgbokwere, Esienumoh, & Uyana, 2015), extramarital affairs (Ojedokun, 2015), and condom use negotiation (Teye-Kwadjo, Kagee, & Swart, 2018). However, after examining the study materials for the extramarital affairs more closely, we decided to drop it because we judged the measure used to be poorly developed. Given our African collaborators' preference for a study on extramarital affairs, we searched for a similarly-themed study and found one in Vera Cruz's (2018) study on partners' willingness to forgive extramarital sexual relationships. Below, we detailed the procedures we followed to select these three studies.

1. Step 1: Study nomination

The study nomination procedure was designed to allow African researchers to nominate studies/effects they consider originally discovered in Africa and of interest to African researchers and populations. To allow African researchers to nominate these studies/effects, we developed an online survey (archived at: <https://osf.io/sgz6x/>) and distributed the survey to a database of 503 email addresses. The database included emails of existing collaborators ($N = 92$) and online-sourced emails of African researchers ($N = 411$), predominantly in psychology departments in educational institutions across 18 African countries. The email addresses were linked to researchers from the following countries: Algeria (83), Cameroon (1), Egypt (12), Ethiopia (2), Ghana (10), Kenya (27), Malawi (37), Morocco (5), Mozambique (3), Namibia (8), Nigeria (30), Rwanda (1), South Africa (233), Tanzania (12), Tunisia (2), Uganda (6), Zambia (24), and Zimbabwe (7). The survey was conducted between February 8 to March 31, 2022.

1.1 Respondents

Out of 503 emails contacted, 152 emails bounced. Out of 351 emails delivered, 41 African researchers responded to the survey, giving a response rate of 11.68%. Twenty-two out of the 41 respondents nominated either broad research areas/topics of interest, and 6 nominated specific studies for replication. The areas of specialization of respondents included general psychology, clinical psychology, work and organizational, health psychology, applied psychology, social psychology, and family psychology. A .csv file of their responses is available at <https://osf.io/fzc92/>.

1.2 Response analysis

In the first part of the survey, we asked respondents to indicate up to 10 areas of research interest in their countries, which led to a total of 113 suggested broad research interest areas/topics. We asked respondents to enter links or other identifying information to articles or studies of previously mentioned research areas/topics of interest in the second part. Studies nominated should have concepts/problems native to their countries or other African countries, conceptualized/defined by an African researcher, and previously investigated in an African country. Ideally, these articles or studies should be well-documented, with a reported p -value $<.01$, pre-registered, and feasible for an online study. Only 2 of 16 articles/studies nominated were considered for the main study evaluation because the other 14 articles either did not use African samples, were theoretical, were not published in a journal, the project leads of ManyLabs Africa have too limited expertise to conduct the study, or the full-texts were not accessible. The studies that are to be included can either be from the published literature or from policy reports.

1.3 Exclusion and ranking of nominated research topics

One-hundred and thirteen research areas/topics were nominated. Because most of these topics were nominated without specific links to specific manuscripts or articles, we had to conduct a more specific search for studies to be replicated. We adopted a 4-step elimination process following our set exclusion criteria to manage the search.

- **Step 1: General-themed topic criterion:** The first step was to exclude general/broad subject areas such as safety, mass media, person perception, or biological psychology.

We excluded 21 out of the 113 research areas/topics based on the general-themed topics exclusion criterion.

- **Step 2: Non-feasible study criterion:** The second step was to exclude research areas/topics for which 1) the lead researchers have too limited research expertise to conduct the study (e.g., communication apprehension and shyness among hyperactive children, climate change on African socio-economic life), 2) the studies are focused on samples that are too hard-to-reach to be easily accessed (e.g., men at risk of suicide; primary school students), and 3) the studies bring up ethical issues that the lead researchers are not equipped to deal with (e.g., sampling victims of sexual violence and traumatic experience). Following these criteria, another 49 nominated research areas/topics were excluded. The final 10 research areas/topics focused on consumerism, dating, marriage, love and hatred, social values and class rivalry, health and security at work, person perception and risk-taking, domestic violence, drugs and societal wellbeing, impacts of child marriage and child pregnancy, gender-based violence, social perception of inter-ethnic marriage, and crime and security.
- **Step 3: No resultant papers/articles criterion:** In Step 3, we searched the [African Journals Online](#) and the [International African institute](#) repositories for potentially published articles in the final 10 research areas/topics mentioned by the nominators. We then used one of the lead authors' (Patrick S. Forscher) connection at the Busara Center for Behavioral Economics to find short research studies that might be suitable, including Busara's "[Off the Record](#)" series. Finally, we searched specific journals in Egypt, Morocco, Tunisia, and South Africa for peer-reviewed African studies. Our search returned 29 articles of interest in these research areas/topics for further assessment.

- **Step 4: Rank papers/articles of interest:** In Step 4, to further streamline the study evaluation process, the study selection lead (a Nigerian researcher) ranked the shortlisted research articles following the preliminary/initial criteria (studies well-documented, reported p -value $<.01$, pre-registered, and feasible for an online study) used for the article search. All ranked and unranked articles/papers were subsequently considered for further evaluation because we had difficulties finding studies that met our set criteria.

1.4 Study search challenges

Finding studies that met all our a priori set criteria was challenging, if not impossible.

Some notable issues that limit our study search and nomination include:

1. Most African studies we found used qualitative methods such as focus groups and interviews, which does not fit well with our predetermined criteria for suitable design and parameters for a large-scale online replication study. Our ideal study should be a quantitative study that adopts a cross-sectional quantitative design of non-clinical samples.
2. Western-developed measures are commonly used in African studies. Because we avoided nominating studies that used Western-developed measures, we had a smaller pool of studies that used African-developed measures or self-developed designs to consider for nomination.
3. Most North African country studies are published in Arabic, a language that the lead researchers have not mastered.

2. Step 2: Evaluation of nominated studies/effects

The second main phase of the study selection involved a comprehensive evaluation of the papers/articles by the research leads using a predetermined three clustered key criteria: locality, feasibility for a multi-site study, and reproducibility. For the *locality criterion*, we considered articles that are 1) potentially of interest and relevant to our African collaborators and 2) adaptable, applicable, and ethically acceptable to the target study populations. For the *feasibility for multi-site study criterion*, we preferred studies that 1) are feasible for online study, 2) use instruments with few items, 3) sampled the general population, and 4) did not rely on an overly complex design. For the *reproducibility criterion*, we favored studies that were 1) preregistered, 2) reported a *p*-value of $<.01$, and 3) were well documented. Although the articles did not meet some criteria, for example, just one of these 29 articles was preregistered, we decided to go forward with six articles. We contacted the corresponding authors of the final nominated articles/papers for research materials such as the original dataset, data analysis script, and actual questionnaire used for the studies. After that, based on the overall study evaluation and materials available, we shortlisted 5 articles (out of these 6 articles) we considered most of interest for the next phase, which is the general study assessment. The shortlisted studies were:

1. Bevan-Dye, A. L., & Akpojivi, U. (2016). South African Generation Y students' self-disclosure on Facebook. *South African Journal of Psychology*, 46(1), 114–129. <https://doi.org/10.1177/0081246315602645>. Hereafter referred to as “self-disclosure study”.
2. Kombo, S. (n.d.). Using behavioural informed communication to drive civic engagement. [Unpublished paper] <https://work.busaracenter.org/case->

[studies/A11%20Community%20Messaging%20&%20Civic%20Participation%20Phase%20%20-%20Uganda.pdf?_ga=2.147651319.1196572791.1657005992-152356769.1647274662](https://doi.org/10.4314/gjpas.v21i2.10). Hereafter referred to as “civic engagement study”.

3. Mgbokwere, D. O., Esienumoh, E. E., & Uyana, D. A. (2015). Perception and attitudes of parents towards teenage pregnancy in a rural community of Cross river state, Nigeria. *Global Journal of Pure and Applied Sciences*, 21(2), 181. <https://doi.org/10.4314/gjpas.v21i2.10>. Hereafter referred to as “teenage pregnancy study”.
4. Ojedokun, I. M. (2015). Extramarital affair as correlate of reproductive health and home instability among couples in Ibadan, Nigeria. *African Journal of Social Work*, 5(2), 1–40. <https://www.ajol.info/index.php/ajsw/article/download/133938/123557>. Hereafter referred to as “extramarital affair study”.
5. Teye-Kwadjo, E., Kagee, A., & Swart, H. (2018). Condom use negotiation among high school adolescents in Ghana: The role of gender. *South African Journal of Psychology*, 48(4), 465–475. <https://doi.org/10.1177/0081246317730282>. Hereafter referred to as “condom use negotiation study”.

The nominated research areas/topics, article exclusion and ranking, evaluation, and shortlisted studies documentation are available at <https://osf.io/v2x5b/>.

3. Step 3: Study assessment

The study assessment was the third and final phase of the study selection. We assessed the 5 shortlisted articles/papers for study interest and relevance (for researcher and country), feasibility for an online study, study adaptability, applicability, and ethics acceptability for the target study populations, and study preference for replication. The purpose was to select the

final three studies that collaborators considered most appropriate for the replication study. The study assessment was conducted between the 14th and 28th of September, 2022.

3.1 Respondents

We surveyed 234 collaborators who signed up for the ManyLabs Africa study. A total of 80 (34%) collaborators participated in the study assessment. A breakdown of the respondents showed that 16 were from Africa (Kenya 1, Malawi 2, Morocco 2, Mozambique 1, Nigeria 8, Rwanda 1, South Africa 1), 41 from Europe (Austria 2, Cyprus 1, Denmark 4, England 1, Germany 9, Ireland 1, Italy 2, Poland 2, Portugal 5, Russia 3, Serbia 3, Slovenia 1, Switzerland 1, The Netherlands 2, Turkey 1, United Kingdom 3), 18 from North America (Canada 3, United States of America 15), and 5 respondents did not indicate their countries. Analysis of the research background revealed that 73 had a psychology background, 2 from medical sciences, and 1 each from economics, education, linguistics, political science, and sociology.

3.2 Method and analysis

We provided respondents with studies/papers' abstracts/summaries and links to the full texts of these studies to allow them to familiarize themselves with each study. We presented questions related to the assessment criteria in varying response scales that best fit each assessment parameter. For study interest, to assess researchers' interest, we asked respondents, "to what extent are you interested in replicating this study?" and to assess their countries' interest, they responded to the question: "to what extent do you think the general population in your country is interested in the topic of this study?" For study relevance, we asked respondents to rate how relevant the studies are to them: "How relevant is this study to you as a researcher?" and to their countries: "How relevant is this study to people in your country?" For study applicability, they responded to the question, "To what extent do you think this effect is

culturally applicable in your country?" For study feasibility, they were asked, "To what extent do you think this study is feasible for an online study in your country?" For ethical acceptability, the question was: "To what extent do you think this study is ethically acceptable in your country?" All parameters mentioned above were measured on a 5-point Likert scale ranging from *not at all* (1; e.g., not at all interested) to *extremely* (5; e.g., extremely interested). For study adaptability, respondents completed an open-ended question: "if the materials are adaptable in their countries and if not, what changes might we require to make to the measures." We asked respondents for additional comments about the study for each study. Finally, for study preference, we asked respondents to rank these studies from the most preferred (scored 1) to the least preferred one (scored 5).

A total of 80 respondents (Africans = 16, Europeans = 41, North Americans = 18, continent not mentioned = 5) responded to these questions. We observed disparities in the number of assessors per study per assessment parameters because, in a few cases, the respondents did not make these assessments. To analyze their responses, we plotted the average ratings for each assessment by studies by respondents' continent and reported the means (M) and standard deviations (SD) scores, as well as the N per rating. We also extracted comments regarding study adaptability and additional comments for each study. The survey, de-identified datasets, and analysis script are available at <https://osf.io/v2x5b/>.

4. Results

4.1 Overall conclusion: Study selection decision

Our findings revealed diverse opinions about the studies, across the three continents, based on the assessment parameters. The final three studies selected are the extramarital affair study, teenage pregnancy study, and the condom use negotiation study. We decided to select

these studies following the evaluation of respondents' assessment of each study vis-a-vis the assessment parameters and the respondents' opinions across the continents. We present the means and standard deviations of these assessment parameters in Figure 3 and Table 7 below.

4.1.1 Study ranking.

First, we report the overall ranking and we did so for all continents together and for the three continents separated. We did so to understand how our collaborators differed in their preference for the studies and to consider whether we should weigh the African contribution more heavily (as we are invested into studying effects that matter to African researchers). We thus considered respondents' study rankings across the continents to judge whether Africans differ in their overall evaluations of these studies compared to their European and North American counterparts. Respondents from Europe and North American had similar studies as their top 3: self-disclosure (first), condom use negotiation (second), and civic engagement (third), while African respondents top 3 were the teenage pregnancy study (First), the extramarital affair study (second), and the self-disclosure study (third). Respondents' impression of these studies thus differed across Europe/North America and Africa. We therefore checked some of the other ratings to make our final decision.

The top 3 ratings polarized along Africa versus Europe/North America for preferred study, study interest, study relevance, and study applicability. Africa consistently rated extramarital affair and teenage pregnancy studies in the top 3 ratings and had inconsistent top 3 ratings for self-disclosure and condom use negotiation. While their European/North American counterparts consistently preferred civil engagement, self-disclosure, and condom use negotiation studies and rated them in the top 3 for all the assessment criteria except for study feasibility, where teenage pregnancy was ranked in the top 3. African and

European/North American respondents were unanimously in their top 3 ratings for study feasibility.

4.1.2 Study interest and relevance.

The dichotomy between North America and Europe on the one hand and Africa on the other hand was again observed in the study interest (i.e., how interesting is this study to you as a researcher?) and relevance (i.e., how relevant is this study to people in your country?) assessments. For European/North American counterparts, the top 3 studies for researchers' interest were: condom use negotiation, self-disclosure, and civic engagement studies, which was thus consistent with their top 3 ranking. The top 3 studies for African researchers' interest are: the extramarital affair study, the teenage pregnancy study, and the condom use negotiation study, which was also consistent with their top 3 ranking.

For North American/European ratings for countries' interest the top 3 were the condom use negotiation study, the self-disclosure study, and the civic engagement study, which was again consistent with the top 3 for the ranking. The top 3 for the African countries' interest were: the teenage pregnancy study, the extramarital affair study, and the self-disclosure study, which was again consistent with their top 3 ranking.

The top three studies for Africa and for Europe/North America study relevance assessments were identical to the study interest assessments. For European/North American respondents the top 3 relevant studies were: civic engagement, condom use negotiation, and self-disclosure, while for their African counterparts the top 3 relevant studies to them were: the extramarital affair study, the teenage pregnancy study, and the condom use negotiation study were of relevance to them.

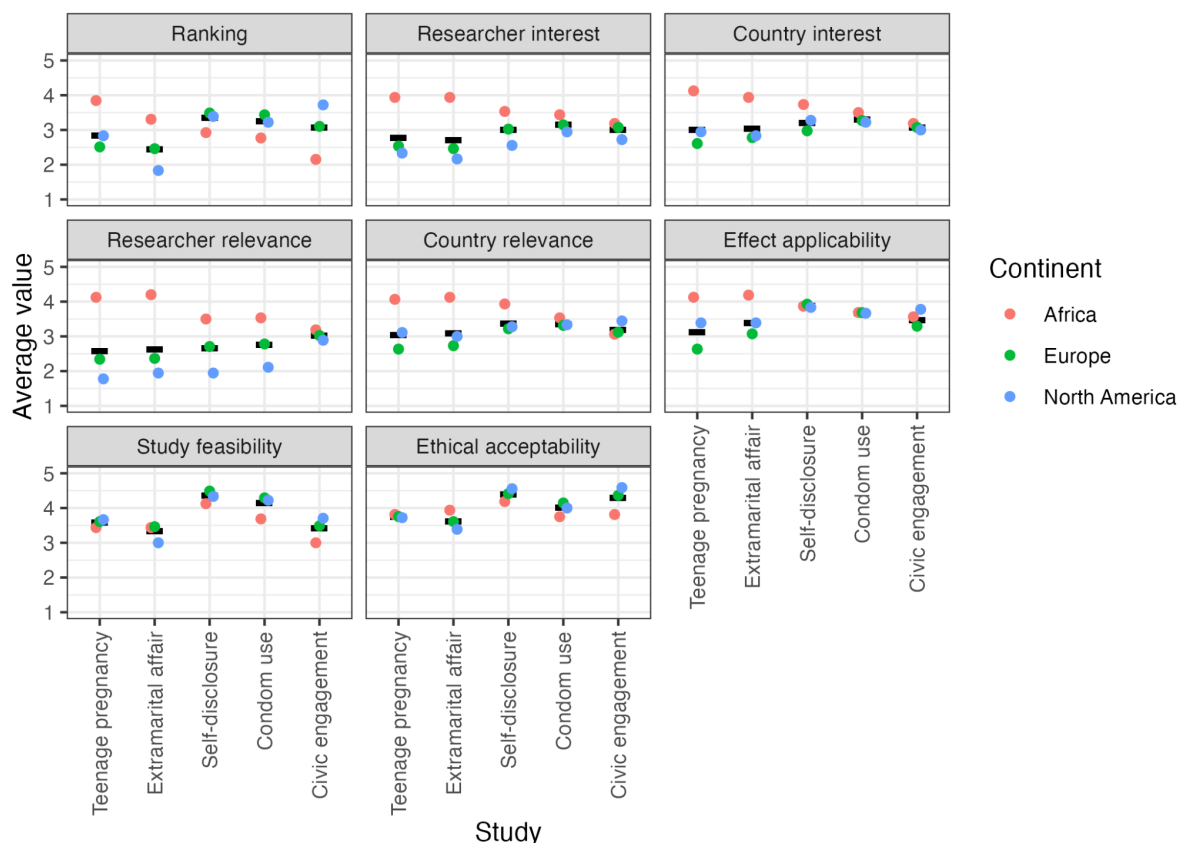
The top 3 relevant studies to European/North American countries were: condom use negotiation study, self-disclosure study, and civic engagement study. While the extramarital

affair study, the teenage pregnancy study, and the self-disclosure study were the relevant studies to African countries. Aside from interest and relevance assessments, it is important to know if these studies are applicable to their local communities to help decide the final studies.

4.1.3 Study applicability.

The study applicability assessment (e.g., to what extent do you think this effect is culturally applicable in your country?) was again consistent with the top 3 studies for interest and relevance and the top-ranked studies for Europe/North America and for Africa. The top 3 for European/North American respondents were: self-disclosure study, condom use negotiation study, and civic engagement study, while the top 3 for African respondents for study applicability were: the extramarital affair study, the teenage pregnancy study, and the self-disclosure study.

Figure 3 Selected study assessment factors per continent and for all regions



Note. On the y-axis are the average values ranging from 1 (least rating) to 5 (highest rating). On the x-axis are the five studies assessed by collaborators. There are 8 panels for the 8 individual assessment parameters. In each panel are dots for each continent (red - Africa, green - Europe, and blue - North America), and also black crossbars for all regions indicating the average value for each continent and for all regions for each study respectively.

Table 7 Selected study assessment parameters’ means and SD for overall and per continent

Parameter	Study	Overall N = 80		Africa N = 16		Europe N = 41		North America N = 18	
		M	SD	M	SD	M	SD	M	SD
Study ranking	Extramarital affair	2.44	1.38	3.31	1.32	2.46	1.31	1.83	1.29
	Teenage pregnancy	2.83	1.22	3.85	1.28	2.51	1.17	2.83	0.92
	Civic engagement	3.11	1.58	2.15	1.07	3.10	1.73	3.72	1.23
	Self-disclosure	3.35	1.43	2.92	1.66	3.49	1.27	3.39	1.61
	Condom use negotiation	3.27	1.29	2.77	1.36	3.44	1.25	3.22	1.31
Researchers’ interest	Extramarital affair	2.64	1.33	3.94	1.06	2.46	1.25	2.17	0.99
	Teenage pregnancy	2.78	1.26	3.94	0.99	2.54	1.07	2.33	1.03
	Civic engagement	2.99	1.25	3.19	0.98	3.07	1.40	2.72	1.07
	Self-disclosure	2.95	1.19	3.53	0.92	3.02	1.27	2.56	0.92
	Condom use negotiation	3.10	1.16	3.44	1.09	3.15	1.17	2.94	1.11
Country interest	Extramarital affair	3.00	1.14	3.94	1.12	2.78	1.11	2.83	0.79
	Teenage pregnancy	2.99	1.10	4.13	0.96	2.61	0.95	2.94	0.80
	Civic engagement	3.05	0.90	3.19	0.91	3.07	0.88	3.00	0.91
	Self-disclosure	3.15	0.99	3.73	0.70	2.98	0.94	3.28	1.07
	Condom use negotiation	3.25	0.98	3.50	1.03	3.27	1.05	3.22	.065
Researchers’ relevance	Extramarital affair	2.59	1.35	4.20	0.77	2.37	1.19	1.94	0.99
	Teenage pregnancy	2.59	1.42	4.13	0.81	2.34	1.26	1.78	0.94
	Civic engagement	2.99	1.19	3.19	0.98	3.02	1.19	2.89	1.32
	Self-disclosure	2.62	1.24	3.50	1.02	2.71	1.23	1.94	0.94
	Condom use negotiation	2.74	1.19	3.53	0.99	2.78	1.11	2.11	1.13
Country relevance	Extramarital affair	3.08	1.10	4.13	0.81	2.73	1.12	3.00	0.59
	Teenage pregnancy	3.04	1.15	4.07	0.88	2.63	1.02	3.11	1.02
	Civic engagement	3.17	0.91	3.06	1.06	3.12	0.81	3.44	0.86
	Self-disclosure	3.33	1.02	3.93	0.79	3.22	1.01	3.28	1.02
	Condom use negotiation	3.33	0.95	3.53	0.92	3.32	1.04	3.33	0.69
Effect applicability	Extramarital affair	3.38	1.11	4.19	0.83	3.07	1.17	3.39	0.69
	Teenage pregnancy	3.15	1.13	4.13	0.50	2.63	1.09	3.39	0.92
	Civic engagement	3.48	0.85	3.56	0.73	3.29	0.78	3.78	1.00
	Self-disclosure	3.82	1.00	3.87	0.64	3.93	0.98	3.83	1.04
	Condom use negotiation	3.66	1.00	3.69	0.95	3.68	1.01	3.67	0.91

Study feasibility	Extramarital affair	3.38	1.22	3.44	1.21	3.46	1.27	3.00	1.19
	Teenage pregnancy	3.60	1.19	3.44	1.09	3.61	1.26	3.67	1.03
	Civic engagement	3.41	1.19	3.00	1.09	3.49	1.19	3.71	1.16
	Self-disclosure	4.34	1.03	4.13	1.03	4.49	0.95	4.33	1.03
	Condom use negotiation	4.16	0.88	3.69	0.87	4.29	0.84	4.22	0.88
Ethics acceptability	Extramarital affair	3.58	0.96	3.94	0.85	3.61	0.95	3.39	0.92
	Teenage pregnancy	3.79	0.99	3.81	0.66	3.76	1.09	3.72	1.02
	Civic engagement	4.26	0.85	3.81	0.75	4.37	0.79	4.59	0.51
	Self-disclosure	4.40	0.81	4.19	0.54	4.41	0.95	4.56	0.62
	Condom use negotiation	4.05	0.83	3.75	0.58	4.15	0.85	4.00	0.97

Note. Five participants did not report their continents. All factors were rated on a 5-point scale of 1 (lowest point) to 5 (highest point).

Thus far, the analysis of the study interest, relevance and applicability have shown that the self-disclosure study and condom use negotiation were consistently rated highly for both African and European/North American respondents. While civic engagement rated highly for European/North American collaborators. In contrast, the African collaborators rated the extramarital affair study and the teenage pregnancy study highly on these parameters. We thus resolved to pay closer attention to studies Africans have rated to be of interest, relevant and applicable. This decision is in line with the objective of the ManyLabs Africa study to investigate African claims and supports our goal of advancing African thought to the European and North American populations (Adetula et al., 2022). Hence, we focused on the study adaptability, feasibility and ethics acceptability assessment for the African-preferred studies: the extramarital affair study, the teenage pregnancy study, condom use negotiation, and the self-disclosure study to determine their suitability for the replication study.

4.1.4 Study adaptability.

Expert judgment of respondents' comments on the adaptability assessment (e.g., are these study measures adaptable in your country?) suggested that 1) Africans are more confident of the adaptability of the extramarital affair study's materials in their countries than European

and North American respondents (see Table 20 in Appendix IV), 2) the teenage pregnancy study and civic engagement study are adaptable to African, European, and North American cultures with minor cultural adaptations or translations (see Tables 21 and 22 in Appendix IV), and 3) the self-disclosure study and condom use negotiation study are very adaptable to the African, European, and North American populations (see Tables 23 and 24 presented in the Appendix IV). To a large extent all these studies are adaptable without any potential consequential adaptability problem indicated by the assessors. Following the study adaptability and previous assessment, the four studies of interest at this stage remained viable for the final selection. Then, we considered these studies feasibility ratings.

4.1.5 Study feasibility.

Respondents from all the three continents had the same studies as their top 3 studies for the feasibility ratings (e.g., to what extent do you think this study is feasible for an online study in your country?). These studies were: the self-disclosure study, the condom use negotiation study, and the teenage pregnancy study with the exception of the North Americans that had the civic engagement study, but not the teenage pregnancy as their top 3. (Table 7; see Figure 3). Again, we were unable to resolve the final 3 studies.

4.1.6 Ethics acceptability.

Finally, when we consider if these studies are ethically acceptable (e.g., to what extent do you think this study is ethically acceptable in your country?). Such as the trend we observed for interest, relevance, and applicability. Respondents' assessments are polarized between Africa and Europe/North America, except that the self-disclosure study was again considered ethically acceptable by all regions. The top three for ethics acceptability for European/North American were: the self-disclosure study, the civic engagement study, and the condom use negotiation study. For their African counterparts were: the self-disclosure study, the

extramarital affair study, and the teenage pregnancy study (see Table 7; Figure 3). Generally, respondents indicated that these studies were ethically acceptable except for few collaborators who commented on specific studies for their sites/countries. For instance, respondents noted that “extramarital affairs are taboo in many countries.” (see Tables 25, 26, 27, 28, and 29 presented in Appendix V for additional comments about these studies).

5. Study selection decision summary

Overall, the study assessment outcomes polarized into Africans versus Europeans/North Americans across studies considered interesting, relevant, and applicable to their populations. Both African and European/North American assessors judged all these studies to be adaptable, feasible for an online study, and ethically acceptable, with few concerns. The extramarital affair and teenage pregnancy stand out as African-preferred studies. In contrast, European/North American assessors preferred the self-disclosure, civil engagement, and condom use negotiation studies.

For the extramarital affair study, African respondents expressed optimism that the study is interesting, relevant, and applicable. However, European and North American respondents were skeptical about its adaptability, applicability, and acceptability (See Table 20 in Appendix IV and Table 25 in Appendix V). However, the extramarital affair study did not report the text of its measures, which left us unable to evaluate the extent to which we could execute the study. We therefore requested the text of the measures from the study authors and found that they did not match the description of the study’s focal variables (e.g., “extramarital affairs affect a couple's sexual relationship” was an item used to measure perceived effects of extramarital affairs on reproductive health). We were also unable to find out which items belong to which subscales, meaning that we would be unable to replicate the study’s findings. We therefore chose not to move forward with this particular study. To respect our African colleagues’

interest in the extramarital affair topic, we searched for studies via Google search databases using extramarital affair articles and names of African countries (one at a time) as search criteria to find a different study on this same topic. We found the Vera Cruz (2018) study on "would Mozambican women really “tolerate” their husbands’ extramarital sexual relationships as socially recommended?" This study uses stronger measurement instruments and has a stronger design. We therefore selected the Vera Cruz (2019) study.

For the teenage pregnancy study, African and, to a lesser extent, European and North American respondents opined that the study is apt and interesting for their communities. However, European/North American respondents expressed concerns about the participants' recruitment challenges due to no access to computers and the internet if we have to recruit poor and rural-dwelling parents for the online study and that they mostly recruit student samples. Some respondents pointed to the need to culturally adapt the study's measure and that the study may not apply to some European countries because of the low rate of teenage pregnancy (for more, see respondents' comments in Table 21 in Appendix IV and Table 26 Appendix V). We selected this study as one of the two most African-preferred studies.

For the civic engagement study, while European/North American respondents rated this study in their top 3 for all the assessment criteria except for study feasibility, Africans didn't rank it top 3 in any of the assessment parameters (see Table 7). African respondents pointed to the research design and the need to adapt the study to online study challenges. Although European and North American respondents expressed minor concerns with the research design, they opined that the study seems applicable to their cultures and that adapting the instrument should be straightforward (see respondents' comments in Table 22 in Appendix IV and Table 27 in Appendix V). Ultimately, we did not select this study because it was not one of the African-preferred studies.

For the self-disclosure study, aside from extramarital affair and teenage pregnancy studies, the self-disclosure was also rated favorably by African assessors for study ranking, continent interest and relevance, applicability, feasibility, and ethics acceptability. European/North American respondents rated it top 3 for all the assessment criteria (see Table 7). Respondents also commented that it is adaptable. However, they were concerned that Facebook, the social media app used in the study, is no longer popular in most countries (See Table 23 in Appendix IV and Table 28 in Appendix V). Although it seems African respondents preferred this study more than the condom use negotiation study, we did not select this study and opted for the condom use negotiation study. The self-disclosure used a structural equation model as its focal results that would be difficult to test in a mixed effects design, and the main interest in the study was on causal claims that are not justified by the study's cross-sectional design.

For condom use negotiation, African respondents rated this study highly for the researcher's interest and relevance and for feasibility. In contrast, European/North American respondents rated it in the top 3 for all the assessment criteria (see Table 7). Respondents considered the study adaptable with minor changes. However, noted respondents' concerns about gender representation and dimension and that participants might resent the foul language used in the study. (See Table 24 in Appendix IV and Table 29 in Appendix V). Besides the respondents' interest in this study, we selected it because it was a good fit for the proposed statistical design.

6. Conclusion

Study selection is a crucial phase of conducting a transnational replication study. Researchers often develop or adopt different approaches for selecting studies for these replications. The ManyLabs Africa study selection approach allows local researchers to be part

of the decision-making process and help reduce selection bias that might influence the choice of these studies. Our selection criteria ensure we have studies that meet local expectations, feasible for online research, are reproducible, and are statistically suitable for testing by-site variability. With these criteria in mind, we selected African-preferred studies: Mgbokwere et al. (2015) teenage pregnancy study, Teye-Kwadjo, et al. (2018) condom use negotiation study, and Vera Cruz (2018) willingness to forgive extramarital sexual relationship study. These studies are sexual/marital relationship-focused that can have significant social relationships, mental health, and medical health implications for Africans and African societies. Although implementing this approach and workflow is time-consuming, it allows for thorough consideration of a broader range of viewpoints in the selection of studies than typifies a top-down approach.

Contrary to some popular big-team science study selection approach where a single or few researchers nominate or propose effects to replicate with little consideration for the priority and applicability to target populations, the ManyLabs Africa study selection adopted an inclusive bottom-up approach to selecting studies. This approach allows investigators to make careful consideration—from the conceptualization stage—to engage the populations where effects were sourced as well as the replication-targeted populations thereby providing a reasonable chance of selecting claims of common interest and relevance to the researchers and to their communities. Evidently, our African and North American/European assessors were divided along study preference underscoring the significance of a robust and inclusive study selection approach to identify relevant research questions for cross-cultural replications across the Global South and Global North divides. This is crucial for broader theory development and applicability of psychological science claims. In Chapter 4, we reported the replication of these African claims and our findings.

Chapter 4: ManyLabs Africa study: Investigating the replicability and generalizability of African effects

In the last decade, various multi-lab multinational replication efforts have emerged to examine the replicability of psychological science (Ebersole, 2016, 2020; Forscher et al., 2022; Klein et al., 2014; 2018, 2022; Open Science Collaboration, 2015). Many of the studied effects in these large replication initiatives originate in the United States and (to some extent) Europe, with little or dubious applicability to other world regions (Adetula et al., 2022). Particularly telling is that out of all the effects replicated in the ManyLabs replication initiatives, zero stemmed from papers by African scholars. Because researchers investigate the topics that are interesting and relevant based on their past experience, the focus on effects discovered in very few countries in these replication efforts – and psychology more generally – leaves open the possibility that the effects discovered by researchers in the United States and Europe will be systematically different than the effects discovered outside these countries. If true, these systematic differences would leave “unknown unknowns” (Adetula et al, 2022) in our knowledge of human psychology, leaving a large swathe of human psychology unexplored. The exclusion of African scholars in the research process is thus not only unethical, it leaves an incomplete view of human psychology.

In an effort to remedy this situation, we can start to focus on African claims to understand what thoughts and theories exist on the continent. To this end, we selected three effects originally discovered in Africa and replicated these among populations of 18 countries across Europe, North America, and Africa. Based on pre-set study selection criteria and our collaborators’ preferences, we selected 1) a study on the perceptions of teenage pregnancy (Mgbokwere et al., 2015), 2) a study on gender differences in attitudes towards condom use

(Teye-Kwadjo et al., 2018), and 3) a study on the willingness to forgive extramarital sexual relationships (Vera Cruz, 2018).

1. The problem with generalizing from North America (and Europe)

Most multi-site replication efforts have focused on effects that originated in the United States. Beyond this narrow focus, it has been well-known that psychological science mostly samples US populations (Arnett, 2008), that white US participants are considered default humans (Cheon et al., 2020), and that psychological science is just WEIRD (Western, Educated, Industrialized, Rich, and Democratic; Henrich et al., 2010). Even as recently as 2018, almost all research published in one of the major journals in psychology, *Psychological Science*, relied mostly on US and Western European samples, while inappropriately generalizing to humans in general (Rad et al., 2018). Despite recommendations to limit the extent to which researchers generalize inappropriately (Simons et al., 2017) and efforts to diversify research samples (Camerer et al., 2018; Coles et al., 2022; Forscher et al., 2022; Klein et al., 2014; 2018; Moshontz et al., 2018; Visser et al., 2022; Zaadnoordijk et al., 2021), little progress has been made for participants outside the rich world to become part of the global research enterprise.

African participants, in particular, have remained undersampled. A recent audit of popular psychological science studies showed that 2,232 of the 3,587 samples used (62%) resided in the US (which only contributes to a little over 4% of the world's population) and just 25 of the 3,587 samples used (less than 1%) resided on the entire continent of Africa (which houses 17% of the world population; Thalmayer et al., 2021).⁴ Even initiatives that put diversity

⁴ Studies varied in participants sampled and Thalmayer et al (2021) only provided information on the full sample, so we don't have a clear view on how many participants took part in different countries.

and inclusion at the core of their mission, like the Psychological Science Accelerator, have largely sampled North American and European populations. In the Psychological Science Accelerator's behavioral impact of COVID-19 studies (Buchanan et al., 2023; Dorison et al., 2022; Psychological Science Accelerator Self-Determination Theory Collaboration, 2022; Wang et al., 2021), for instance, only 5,950 African participants (12.5% of all participants) took part.

It is not only African participants that are underrepresented; African researchers – and thus ideas – are similarly underrepresented. Thalmayer et al. (2021) found that 9,660 out of 15,085 (64% of first authors and co-authors) of research published in six popular APA psychology journals between 2014-2018 worked in the United States. Only 27 of these authors worked in Africa (less than 1%). In the Psychological Science Accelerator's COVID-19 studies, 13 of the entire author list (2.7% of all authors) were affiliated with African institutes, and to date, no African author has led a study with the Psychological Science Accelerator. Such research collaborations between the Global North and Africa and other big-team science investigations often neglect or fail to meet Africans' expectations, interests, and priorities (Adetula et al., 2022; Mpofo, 2002; Mughogho et al., 2023; Pillay, 2017; Zeineddine et al., 2021). One of the Psychological Science Accelerator's projects, for example, focused on the well-known trolley dilemma (Bago et al., 2022). Had African researchers designed and selected a study on morality, would they have focused on fictional trains running over fictional people? Or would they have instead preferred studying the Yoruban personhood concept of *Omoluabi* (a person who respects honor, works hard, and gives to the community) or the South African concept of *Ubuntu* (the idea of having a moral and communal obligation to others, based on the belief that a universal bond of sharing connects all humanity; Adetula et al., 2022)?

Without a serious effort to engage scholars from diverse geographical regions, we don't know. The risk remains that psychological science will miss out on variations that are crucial

to understanding different peoples. If they continue to focus on Western-originated ideas, multi-lab replication efforts could then simply supercharge the existing US dominance, with ever more resources going to already rich researchers (Adetula et al., 2022; Dutra, 2021; IJzerman et al., 2021). As a result, theories remain non-generalizable, incomplete, or worse, just false. One way to start changing this state of affairs is by explicitly focusing on research generated outside the United States and Europe for replication efforts. We, therefore, made a modest first step to changing this state of affairs by involving African researchers in selecting and replicating three claims discovered in three African countries in a multi-site study in 18 countries.

2. Selected studies for our replication

In order to select studies originally discovered and relevant to African populations, we first consulted with experts working with populations in the Global South to determine our study selection approach detailed earlier in Chapter 3a. Following this consultation, we asked African researchers to nominate research questions. Based on the themes nominated by our African colleagues, we searched, evaluated, and shortlisted five studies, which our (North American, European, and African) collaborators then evaluated. As preferences between North American and European collaborators, on the one hand, and African collaborators, on the other hand, differed, we decided to follow our African collaborators' preferences (reported in Chapter 3b). The studies we selected for replication were:

Mgbokwere, Esienumoh, and Uyana (2015) - teenage pregnancy study. This study investigated the effect of 309 Nigerian parents' (aged 35 and above; from Akpabuyo Local Government Area) socio-demographic attributes on perception and attitudes towards teenage pregnancy. They found that the effects of parents' gender, religion, and level of education were significant, while parents' age was not significant. However, we did not consider these socio-

demographic effects in our replication study because we collected convenience samples. Our data was not representative of the given demographics. We instead replicated the percentage agreements of Nigerian parents' attitudes towards pregnancy in teenage girls; attitudes were mixed (with negative, supportive, and positive attitudes). We provide the original statements, percentage agreement, and the percentage confidence intervals in Table 8. Our focal effects were the agreement proportion intercepts for the 13 items/statements across Europe, and Africa, as well as the differences between Africa and Europe. We intended to include North America in the analysis and conduct the North America versus Africa comparison but could not because the North American sample had very few parents over the age of 35, so we focused on Europe vs Africa.

Table 8 Mgbokwere, Esienumoh, and Uyana (2015) study's frequencies, percentages, and our calculated percentage CI

Perception of teenage pregnancy			
Item	Agree	Disagree	%CI for Agreement
A teenage pregnant girl is a deviant	288 (93%)	21 (7%)	90.4% - 96%
A teenage pregnant girl does not have a future	259 (84%)	50 (16%)	79.7% - 87.9%
Teenage pregnancy makes parents devastated	265 (85.8%)	44 (14.2%)	81.9% - 89.7%
Teenage pregnancy means that parents have failed	264 (85%)	45 (15%)	81.5% - 89.4%
Teenage pregnancy results in stigmatization	276 (89%)	33 (11%)	85.9% - 92.8%
Teenage pregnancy is the result of western education	217 (70.2%)	92 (29.8%)	65.1% - 75.3%
Attitudes towards teenage pregnancy			
Item	Agree	Disagree	%CI for Agreement
Drive the girl away from the family in anger	101 (32.7%)	208 (67.3%)	27.5% - 37.9%
Stop her schooling	154 (49.8%)	155 (50.2%)	44.2% - 55.4%
Give her out in early marriage	194 (62.8%)	115 (37.2%)	57.4% - 68.2%
Abort pregnancy	49 (15.9%)	260 (84.1%)	11.8% - 19.9%
Send her to mission home for care	215 (69.6%)	94 (30.4%)	64.4% - 74.7%
Support girl to have the baby	178 (57.6%)	131 (42.4%)	52.1% - 63.1%
Force the person responsible for the pregnancy to accept the responsibility of her care	260 (84%)	49 (16%)	80.1% - 88.2%

Note. N = 309. Frequencies and percentages of Mgbokwere, Esienumoh, and Uyana (2015) study's 13 items/statements measuring parents' perception and attitudes towards teenage pregnancy and our calculated confidence intervals⁵ for the agreement proportions

Teye-Kwadjo, Kagee, and Swart (2018) - condom use negotiation study. The authors investigated gender differences in 5 attitudinal variables toward condom use among 684 Ghanaian adolescents (aged 14-20; 335 males and 349 females). Despite that, again, as we relied on convenience samples, we replicated the finding that male Ghanaian adolescents held more positive attitudes towards condom use, greater feelings of control over condom use, a higher number of self-reported condom-protected sexual behaviors than their female counterparts (given the lack of representative sampling, we took care in limiting our conclusions regarding gender). Our focal variables were the statistically significant findings on these positive attitudes. There were no significant gender differences for subjective norms regarding condom use and intentions towards future condom use, which we, therefore, did not include in our replication.

Vera Cruz (2018) - extramarital sexual relationship study. In this study, 225 Mozambican women (aged 18 - 64, all from Maputo and Matola, southern Mozambique) responded to 24 scenarios at four levels of a husband's emotional involvement in an extramarital sexual relationship (hereafter referred to as emotional involvement; low, low-moderate, moderate-high, and high), at three levels of regret expressed by a husband (hereafter referred to as regret expressed; low, intermediate, and high), and two levels of a husband's commitment to marriage (hereafter referred to as commitment to marriage; low and high) and reported to what extent they are willing to forgive the cheater on an 11-point scale. The author reported that two out of three main effects, and two out of four interactions were significant. For these findings, 1) women reported more willingness to forgive a husband involved in a less emotionally involved

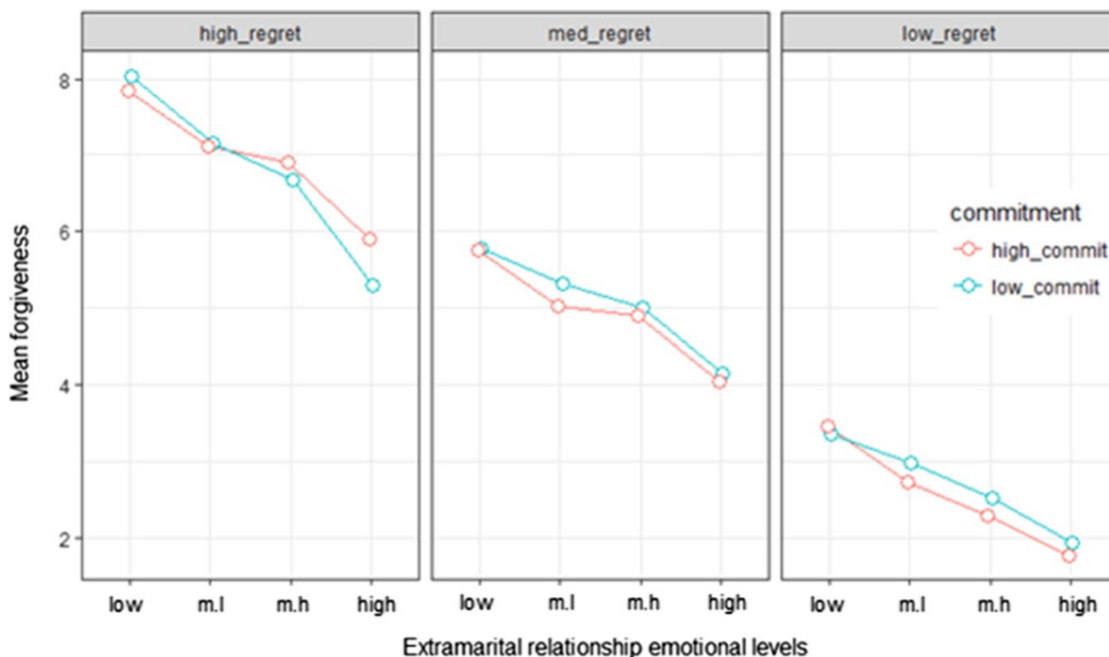
⁵ We calculated the confidence intervals ourselves using the agreement frequencies reported in the original study. The analysis script for the CIs calculations is available at <https://osf.io/mqejn>.

extramarital affair than a more involved affair. Also, 2) women were more willing to forgive a husband that expressed more regret than a husband that expressed less regret. The interaction between emotional involvement and the expression of regret was significant, 3) such that when expressed regret was low, women were more willing to forgive a husband who had a low emotional involvement compared to a husband who had a high degree of involvement. By contrast, when expressed regret was high, women were more willing to forgive a husband who had a low involvement than high involvement. The interaction between emotional involvement and commitment to marriage was significant, 4) when commitment to marriage was low, women were more willing to forgive a husband who had low involvement than high involvement. By contrast, when commitment to marriage was high, they were less willing to forgive a husband who had a high involvement than low involvement (see Figure 4). Although the interaction effects were significant, they were small in magnitude, meaning they may not be robust in replication. We, therefore, focus on the main effects of emotional involvement and regret expressed on willingness to forgive. Notwithstanding, we examined the main effect of commitment to marriage and the interactions between the three variables for potential cultural differences across the target regions.⁶ Our fictional plots of the expected main and interaction effects are available at <https://osf.io/94ebq>. For feasibility reasons⁷ and after having reached an agreement with the original author, our design consists of two levels of emotional involvement (low versus high), at two levels of regret expressed (low versus high), with two levels of commitment to marriage (low versus high).

⁶ Although our focal effects were the main effects of emotional involvement and regret expressed, our mixed effects models tested for all the main effects including for commitment to marriage, Africa versus Europe, and Africa versus North America as well as all their interactions.

⁷ For feasibility purposes, we could not use the original study full design of 4x3x2 ANOVA design, four degrees (low, low-moderate, moderate-high, and high) of emotional involvement at three levels of regret expressed (low, intermediate, and high) expressed, and two degrees of commitment to marriage (low and high), we opted for 2x2x2 design two levels of low and high for emotional involvement, for regret expressed, and for commitment to marriage.

Figure 4 Willingness to forgive mean ratings as an operation of the three factors (sourced from Vera Cruz, 2018).



Note. Participants’ willingness to forgive mean ratings 1) are higher for low and high emotional levels if the regret expressed is high compared to when the low regret condition, and 2) differ with the commitment levels involved in the couple relationship/marriage. There are three panels for the three different levels of regret (high, moderate, low). For each panel, the participants’ willingness to forgive mean ratings are on the vertical axis, the extramarital relationship emotional levels (low, moderate-low, moderate-high, high) on the horizontal axis, and the commitment levels (low, high) are the curves within the panels. From “Would Mozambican women really “tolerate” their husbands’ extramarital sexual relationships as socially recommended?” (Vera Cruz, 2018). Reprinted with permission from the author.

For all effects, we first tested how well they replicated across regions, then whether our African and North American participants differed in their findings, then whether our African and European participants differed in their findings. Finally, we examined heterogeneity across our entire sample and within the three regions.

3. Replication Hypotheses

In the current study, we initially examined whether the focal effects for the current replication held first among the subgroups sampled in the original studies and then the overall sample. We then examined whether our participants in Africa and North America differed for

each hypothesis, followed by whether our participants in Africa and Europe differed for each hypothesis (see Table 9).

As the teenage pregnancy study was descriptive, we only examined whether the confidence interval of the replication study overlapped with the original study. For the condom use negotiation study, since we are just testing if the effects are significantly different from zero, we considered whether the p -value for the test of each replication effect was less than a pre-specified alpha: .0167⁸. For the willingness to forgive extramarital sexual relationships study, we removed several levels, so the confidence interval would not be comparable. Hence, our replication criteria test is to determine whether or not the current study focal effects' p -values are lower than the preset alpha .0005⁹.

⁸ The pre-specified p -value (0.0167) is the average of the original adjusted p -values for the t -test for attitudes towards condom use (p -value = .006, partial η^2 = .011), perceived control over condom use (p -value = <.001, partial η^2 = .019), normative use (p -value = .479, partial η^2 = .001), intention to use condom (p -value = .024, partial η^2 = .007), and self-reported condom use (p -value = <.001, partial η^2 = .042).

In addition to this p -value replication test, we initially planned a replication test to compare the focal effect confidence intervals to a preset focal effect size (-.11) which is the difference of the original mean differences between males' and females' attitudes toward condom use (.23) and perceived control of condom use (.34). However, we jettisoned this replication test after we had preregistered the study and relied on the p -value replication test solely because we didn't thoroughly consider or measure a quantity test on the difference of the mean differences of the variables.

⁹ The preset p -value (0.0005) is the average of the original p -values for the F -test for the emotional involvement (p -value = .000), and regret expressed (p -value = .001) interactions.

Table 9 Replication hypotheses

Study	Original hypothesis targeted for replication	Statistical models using lme4 syntax	Focal test hypothesis	Secondary tests	
				Africa v EU	Africa v US
Teenage pregnancy study	*A pregnant teenage girl is a deviant	$agree_v_disagree \sim 1 + africa_v_eu_c + (1 site)^{10}$	$b_{intercept(prob.)} = 0.149$	$b_{africa_v_eu_c} = 10.73$	-
	A pregnant teenage girl do not have future	Same models as above	$b_{intercept(prob.)} = 0.067$	$b_{africa_v_eu_c} = 4.58$	-
	Teenage pregnancy makes parents devastated	Same models as above	$b_{intercept(prob.)} = 0.550$	$b_{africa_v_eu_c} = 16.83$	-
	Teenage pregnancy means that parents have failed	Same models as above	$b_{intercept(prob.)} = 0.155$	$b_{africa_v_eu_c} = 2.68$	-
	Teenage pregnancy results in stigmatization	Same models as above	$b_{intercept(prob.)} = 0.719$	$b_{africa_v_eu_c} = 2.14$	-
	Teenage pregnancy is the result of western education	Same models as above	$b_{intercept(prob.)} = 0.132$	$b_{africa_v_eu_c} = 3.49$	-
	Parents should drive girl away from the family in anger	Same models as above	$b_{intercept(prob.)} = 0.017$	$b_{africa_v_eu_c} = 11.38$	-
	Parents should stop her schooling	Same models as above	$b_{intercept(prob.)} = 0.023$	$b_{africa_v_eu_c} = 7.59$	-
	Parents should give her out in early marriage	Same models as above	$b_{intercept(prob.)} = 0.053$	$b_{africa_v_eu_c} = 7.20$	-
	Parents should abort pregnancy	Same models as above	$b_{intercept(prob.)} = 0.121$	$b_{africa_v_eu_c} = 1.16$	-
	Parents should send her to a maternity group home for pregnant teenage girls for care	Same models as above	$b_{intercept(prob.)} = 0.137$	$b_{africa_v_eu_c} = 10.67$	-
	Parents should support girl to have the baby	Same models as above	$b_{intercept(prob.)} = 0.839$	$b_{africa_v_eu_c} = 0.91$	-
Parents should force the man/boy responsible for the pregnancy to accept responsibility of her care	Same models as above	$b_{intercept(prob.)} = 0.421$	$b_{africa_v_eu_c} = 5.11$	-	

¹⁰ We revised the original two models ($agree_v_disagree \sim 1 + c1 + c2 + (1 | site)$, and $agree_v_disagree \sim 1 + africa_v_eu + africa_v_us + (1 | site)$) for the teenage pregnancy focal test analyses after preregistering the study because we had to exclude the North America sample for the subgroup analysis as the North American subset sample of parents aged 35 and above is too small and cannot be comparable to the African and European subgroups for this demographic. We, however, used the original models for the full sample analyses that included the North American sample large enough for the comparison test (see results in Appendix VI, Table 31).

Condom use negotiation study	Male college students aged 18-20 will have more positive attitudes towards condom use than their female counterparts.	attitude $\sim 1 + \text{gender} * (c1 + c2) + (1 + \text{gender} \text{site})$	$b_{\text{gender}} = .01$	-	-
		attitude $\sim 1 + \text{gender} * (\text{africa_v_eu} + \text{africa_v_us}) + (1 + \text{gender} \text{site})$	-	$b_{\text{gender}*\text{africa_v_eu}} = .04$	$b_{\text{gender}*\text{africa_v_us}} = .11$
	Male college students aged 18-20 will have greater feelings of control over condom use than their female counterparts.	control $\sim 1 + \text{gender} * (c1 + c2) + (1 + \text{gender} \text{site})$	$b_{\text{gender}} = -15$	-	-
		control $\sim 1 + \text{gender} * (\text{africa_v_eu} + \text{africa_v_us}) + (1 + \text{gender} \text{site})$	-	$b_{\text{gender}*\text{africa_v_eu}} = .34$	$b_{\text{gender}*\text{africa_v_us}} = .33$
	Male college students aged 18 - 20 will have a higher number of self-reported condom-protected sexual behaviors than their female counterparts.	reported $\sim 1 + \text{gender} * (c1 + c2) + (1 + \text{gender} \text{site})$	$b_{\text{gender}} = -.48$	-	-
		reported $\sim 1 + \text{gender} * (\text{africa_v_eu} + \text{africa_v_us}) + (1 + \text{gender} \text{site})$	-	$b_{\text{gender}*\text{africa_v_eu}} = .87$	$b_{\text{gender}*\text{africa_v_us}} = .55$
Extramarital sexual relationship study	When the emotional involvement is low, heterosexual women are significantly more willing to forgive than when the emotional involvement is high.	forgive $\sim \text{regret} * \text{commitment} * \text{emotion} * (c1 + c2) + (\text{regret} + \text{commitment} + \text{emotion} \text{ID}) + (\text{regret} + \text{commitment} + \text{emotion} \text{site})$	$b_{\text{emotion}} = -.21$	-	-
		forgive $\sim \text{regret} * \text{commitment} * \text{emotion} * (\text{africa_v_eu} + \text{africa_v_us}) + (\text{regret} + \text{commitment} + \text{emotion} \text{ID}) +$	-	$b_{\text{emotion}*\text{africa_v_eu}} = .20$	$b_{\text{emotion}*\text{africa_v_us}} = -.14$

(regret + commitment + emotion | site)

When the regret expressed is high, heterosexual women are more willing to forgive compared to when the regret expressed is low.

Same models as above

$$b_{regret} = 1.17$$

$$b_{regret**africa_v_eu} = 1.25 \quad - \quad b_{regret**africa_v_u} = -1.34$$

Note. EU – Europe, US – North America

A summary table of the current study hypotheses, statistical effect, and the focal effects and by-regions' proposed results.

*We did not target the original hypothesis for the replication test because we could not ascertain how the authors quantified participants' responses and determined the mean values

4. Method

4.1 Participants, Power analysis, Recruitment, and Ethics

We attempted a priori power analyses, but there were too many unknowns to successfully do so¹¹. We were, for instance, unable to make reasonable assumptions for the by-site and by-subject variance parameters. As a result, we decided to simply recruit a large number of sites (Africa = 109; Europe = 108; North America = 44) and to set a minimum recruitment target of 80 participants per site. We aimed to collect data from at least 30 sites (10 per target continent).

To get a sense of the precision we achieved, we took a different approach for the teenage pregnancy study on the one hand and the condom use negotiation and extramarital sex studies on the other. For the condom use negotiation and extramarital sex studies, we conducted a sensitivity power analysis after data collection to explore the effect sizes that we could detect with our design, given the sample size we achieved and a particular level of desired power (Lakens, 2022). We conducted these power analyses using the *simr* package, testing a sequence of effect sizes from small to large for each study's focal effect. These analyses revealed that, for the condom use negotiation study, our achieved sample was sufficient to detect with at least 90% power a gender difference in self-reported condom use $-.5$ (compared to $-.74$ reported by the original authors), and difference in forgiveness between low and high regret conditions of $.6$ (compared to 4.21 reported by the original authors).

Because the teenage pregnancy study was exploratory and did not involve testing a particular hypothesis, we simply drew 1000 bootstrap samples and computed confidence

¹¹ We intended to reanalyze the original studies to establish if the original findings were as robust as reported in these studies (Adetula, et al., 2021). However, as we don't have access to the original datasets, we were unable to conduct a computational reproducibility check.

intervals around the estimated intercept, allowing us to assess how the width of these confidence intervals varied given the sample size we achieved. We drew our bootstrap samples using the *lmeresampler* package and found that, across the samples, the average confidence interval had a width of .81 log-odds units, and just 13.4% of the confidence intervals had a width exceeding 1 log-odds unit. Overall, the results of our sensitivity power analyses and bootstrapping analysis suggest that our study was sufficiently precise.

Sites were responsible for sourcing and recruiting participants, distributing the survey links, and collecting data locally. We provided them with anonymous Qualtrics links (English and/or translated versions) to collect data from their local communities. To recruit participants, contributing labs invited participants from their community to take the survey either on their smartphone, their own computer or iPad, or on a computer in a computer lab (if the collaborator had it available). In turn, we asked participants how they were recruited to account for the recruitment source. After we excluded 1352 out of the 7361 participants for missing data¹², 43 sites (Africa = 16, Europe = 17, North America = 10) recruited participants at a university/institution [$N=4185$], via a friend or colleague [$N=335$], via online participant recruitment sites: Amazon mTurk [$N=5$], Prolific Academic [$N=518$], or social media platforms such as WhatsApp [$N=206$], Facebook [$N=207$], X (formerly Twitter) [$N=7$], or Mastodon [$N=1$], unreported [$N=230$], or another recruitment source [$N=315$]).

A total of 6009 participated in our online survey from the 43 sites across Africa (1136), Europe (3012), and North America (1861). Participant populations varied across three studies (See Table 10 for the current study's population and sample characteristics deviations from the original studies). To make sure the study was feasible to conduct in a package and that it met ethical standards in all participating countries, we sampled participants who were 1) of adult age (i.e., at least 18 years of age), 2) in a relationship or not, and 3) parents and non-parents for

¹² Participants with less than 40% completion rate

all three studies. We then analyzed subsets of participants to ensure equivalence from original to replication studies.

We further excluded participants who spent less than 343 seconds completing the survey¹³ ($N=203$), which gave us an exclusion rate of 3.4%. The remaining 5806 final samples used for the analysis comprised 1110 participants in Africa, 2911 participants in Europe, and 1785 participants in North America. See Table 11 for the number of participants per country and Table 12 for participants' demographics per continent.

¹³ Participants who spent less than 1/3 lower than the overall median time spent completing the survey

Table 10 Original and replication samples description

Study	Original sample	Focal replication sample	Deviation justification
Teenage pregnancy	Parents aged 35 years and above	Parents aged 35 years and above	As we collected three samples at once, the most pragmatic solution was to enlarge the sample to participants over 18. By enlarging our sample, we could further test the generalizability of the claim to the focal sample parents under the age of 35 as well as to the entire participants, non-parents and parents aged 18 years above.
Condom use negotiation	Adolescents college students aged 14 to 20 years	College students 18 to 20 years old ¹⁴	Our bundling meant the most pragmatic solution was to enlarge the sample and to collect participants above 20 and non-college students. However, we also deviated as we did not include participants under 18, because of 1) the sensitive nature of the subject, 2) most countries' legal responsibility age is 18 years, and 3) obtaining ethics approval and consent for persons under 18 years would be challenging. By enlarging our sample, we could test the generalizability of the claim to our focal sample of college students 18 to 20 years as well as to the entire participants, college and non-college students aged 18 years and above.
Extramarital Sex study	Mozambican women aged 18 - 64	Heterosexual women aged 18 years and above	Again, our bundling meant the most pragmatic solution was to enlarge the sample and to collect data from people other than women. This deviation allowed us to generalize to first heterosexual women aged 18 years and above as well as to other populations (e.g., male and non-heterosexual females).

Note. A summary table showing the study population and sample characteristics deviations from the original studies and justification for the deviation

¹⁴We excluded 1106/5806 (19%) non-heterosexual male and female participants for this study because 1) the original study focused on heterosexual male-female differences and 2) non-heterosexual genders are hard to reach.

Table 11 Number of participants per country

Region	Country	Number of participants
Africa	Burkina Faso	9
	Gabon	186
	Morocco	304
	Nigeria	591
	Tanzania	20
		1110
Europe	Bulgaria	65
	Germany	309
	Ireland	101
	Italy	89
	Netherlands	82
	Poland	200
	Portugal	219
	Serbia	356
	Slovenia	148
	Turkey	185
	United Kingdom	1157
		2911
North America	Canada	128
	United States	1657
		1785
Overall		5806

Table 12 Participants' demographics

Demographic	Category	Overall	Africa	Europe	North America
Gender	Male	1492	358	728	406
	Female	4025	634	2062	1329
	Agender	21	-	15	6
	Genderfluid	28	-	14	14
	Genderqueer	15	-	10	5
	Prefer to self-describe	35	-	21	14
	Prefer not to say	63	39	14	10
	Unreported	127	79	47	1
Sexual orientation	Heterosexual	3590	-	2272	1318
	Homosexual	198	-	114	84
	Bisexual	649	-	366	283
	Prefer to self-describe	135	-	77	58
	Prefer not to say	75	-	35	40
	Unreported ¹⁵	1159	1110	47	2

¹⁵ We posed the sexual orientation question only to participants in countries where sexual orientation identifications and relationships are legally accepted.

Age	<i>M</i> years (<i>SD</i>)	26.6(10.6)	27.2(8.2)	30.4(12.2)	20.1 (3.9)
Marital status	Not married and not in a relationship	2167	431	870	866
	Not married but in a relationship	1898	308	831	759
	Living with a partner	572	30	470	72
	Married	733	160	531	42
	Divorced	80	21	52	7
	Separated	61	9	45	7
	Widowed	14	3	11	0
	Prefer not to say	148	64	53	31
	Unreported	133	84	48	1
For participants currently in a relationship: Years in current relationship	<i>M</i> months (<i>SD</i>)	63.0 (92.7)	46.2 (75.4)	85.4 (107.7)	25.5 (36.0)
If you are in a sexual relationship, do you typically use condoms while having intercourse?	No	1730	200	1140	390
	Yes	1163	181	615	367
	Prefer not to say	296	107	73	116
	Unreported	2617	622	1083	912
Aside from condoms, do you use other forms of birth control?	No, only condom	783	118	461	204
	Yes, not only condom	358	61	145	152
	Prefer not to say	22	2	9	11
	Unreported	0	0	0	0
Parental status	No	4534	747	2069	1718
	Yes	1083	243	778	62
	Prefer not to say	45	25	16	4
	Unreported	144	95	48	1
Formal education years	<i>M</i> years spent (<i>SD</i>)	15.5(3.7)	17.1(5.3)	15.4(3.7)	14.8(1.9)
Studenthood status	Not a student	887	174	710	3
	Undergraduate student	4256	651	1842	1763
	Postgraduate student	512	189	305	18
	Unreported	151	96	54	1
Religion	Buddhism	65	0	17	48
	Caodism	1	0	0	1
	Christianity	2268	608	931	729
	Confucianism				

Hinduism	47	0	12	35
Islam	489	290	167	32
Judaism	36	0	5	31
Shinto	0	0	0	0
Sikhism	9	0	3	6
Taoism	2	0	1	1
Traditional African religion	6	1	2	3
Other (please specify)	180	13	56	111
I am not religious but am agnostic	601	14	379	208
I am not religious but am spiritual	739	35	469	235
I am not religious but am atheist	845	8	669	168
Prefer not to say	362	43	146	173
Unreported	156	98	54	4

Note. *SD* - standard deviation,

The frequencies of participants’ gender, sexual orientation, their average and standard deviation age, marital status, use of condoms during intercourse, use of other birth control, *M* year, and *SD* in their current relationships, parental status, *M* and *SD* of years spent acquiring formal education, studenthood status in a university/college, and religious affiliation for entire participants and per continent.

Each site was responsible for its own ethics approval. While twenty-one sites obtained approval from their local institutions, eighteen sites adopted an Ashland University-approved protocol. One site indicated that their institution did not require approval for the type of studies we conducted, and five presented exemption/permission letters. The style and the amount of compensation varied with local conventions. More information ethics can be found at <https://osf.io/9fhu5/>.

5. Procedure

The data were centrally collected and collated by the research lead at LIP/PC2S via a Qualtrics subscription managed by SCREEN at Université Grenoble Alpes, France. Data collection was conducted between November 21, 2023 to April 8, 2024. Contributing labs provided the anonymous Qualtrics links to potential participants. We developed an online questionnaire containing items used in the original studies, a measure to assess sociosexual

orientation (Penke & Asendorpf, 2008), and demographic questions. We created two versions of our survey, one including an inclusive way to ask about gender and sexuality, and one with a less inclusive way to ask about gender and which did not ask about sexuality, for those countries in which non-heterosexuality is forbidden. We did so to minimize the risk for our participants. The three replication studies were offered in a random order. Items within question blocks were also randomized. Once they completed that portion of the study, they responded to 14 demographic questions, and lastly, they were debriefed. It took 17 minutes on average for participants to complete this survey. The final questionnaire in Qualtrics and as Word export is available via <https://osf.io/g4u73/>.^{16,17}

On our OSF page (<https://osf.io/52mrt/>), we documented study-per-study changes made to the original studies in order to further improve statistical design and analysis, participants' inclusion, study feasibility, and measures' meaningfulness and relevance.

6. Measures

6.1 Teenage pregnancy study measures.

The Teenage Pregnancy Study's objective is to assess the attitude of parents towards pregnant teenage girls. To measure this, participants were asked to respond on a 2-point Likert scale whether they agreed or disagreed to 1) a 6-item measure on respondents' perception of the pregnant teenage girl and teenage pregnancy (e.g., "A teenage pregnant girl does not have a future"), and 2) a 7-item measure of perceived/expected parents' behavior towards the pregnant girl child (e.g., "give her out in early marriage"). The original authors report a test-

¹⁶ In the original extramarital sex measure, only female participants answered the scenarios. In our study, both male and female participants who responded to being (mostly) heterosexual or (mostly) homosexual responded to the same original scenarios.

¹⁷ As in the original willingness to forgive extramarital sexual relationship study, we presented 2 scenarios for familiarization purpose: 1) the scenario of high commitment x moderate low emotion, high regret, and 2) a scenario of low commitment x moderate low emotion x low regret), Unlike the original study, we did not request participants to review their ratings at the familiarization stage; instead, we implemented the Qualtrics required response validation to ensure participant response to the familiarization scenarios.

retest reliability coefficient of 0.90. For this present study, we used all the 13 items. The full items of the measure are available at <https://osf.io/a52eb>.

6.2 Condom use negotiation study measure

This study examined gender differences in 1) attitudes, 2) perceived control, and 3) actual use of condoms. To measure attitudes towards condom use, the original authors used a 5-item that measures attitudes towards condom use (Basen-Engquist et al., 1999; Carmack & Lewis-Moss, 2009; e.g., “I am in favor of using condoms”) on a 7-point Likert scale ranging from Strongly Disagree scored 1 to Strongly Agree scored 7. They reported an Alpha coefficient (α) of 0.69. For this present study, we used all the 5 items.

To assess perceived behavioral control over condom use, they used a 4-item that measures participants’ perceived difficulty and efficacy of condom use (Carmack & Lewis-Moss, 2009; Jemmott et al., 2007; e.g., I can use a condom correctly)¹⁸ on a 7-point Likert scale ranging from Strongly Disagree scored 1 to Strongly Agree scored 7. They reported an Alpha coefficient (α) of 0.75. For this present study, we used all the 4 items.

To assess participants’ actual condoms use behavior, they used a 7-item measure of self-report of condom use behavior in the past 3 months (Holland & French, 2012; Walsh et al., 2011; e.g., how often have you had sex with your regular boyfriend/girlfriend with a condom in the past 3 months)¹⁹ on a 7-point Likert scale ranging from Never scored 1 to Very often scored 7. All items were scored directly except for Item 4 (How often did you refuse to have sex with a partner because they would not use a condom in the past 3 months?) scored

¹⁸ Teye-Kwadjo, Kagee, and Swart (2018) slightly modified one of the items for meaningfulness and for contextual relevance.

¹⁹ Teye-Kwadjo, Kagee, and Swart (2018) slightly modified two items for meaningfulness and for contextual relevance.

inversely. They reported an Alpha coefficient (α) of 0.89. For this present study, we used all the 7 items. The full items of these measures are available at <https://osf.io/275wc>.

6.3 Extramarital sex study measure

The goal of the Extramarital Sex Study was to assess the women's willingness to forgive extramarital sexual relationships. To measure this, participants read 24 scenarios following a 4x3x2 design of 4 levels (low, low-moderate, moderate-high, and high) of husband's emotional involvement in an extramarital sexual relationship, 3 levels (low, intermediate, and high) of degree of regret expressed by the husband, and 2 levels (low and high) of husband's commitment to marriage. Respondents marked/picked a point to indicate their degree of willingness to forgive on an 11-point scale ranging from 0 ("certainly would not") to 10 ("certainly would"). For feasibility reasons, we reduced this to a 2x2x2 design using only the scenarios for low and high levels for the three factors. The full items of these measures are available at <https://osf.io/728zw>.

6.4 General measures

Sociosexual orientation (Penke 2010). We included sociosexual orientation, an inclination to engage in uncommitted sexual relations (Hone et al., 2020; Penke & Asendorpf, 2008; Simpson & Gangestad, 1991), as an exploratory measure as a potential way to explain differences in replication success across samples. After all, 1) all original studies we replicated focused on sexual/interpersonal relationships and 2) one prior replication study was able to explain non-replication with this measure (see IJzerman et al., 2014). This instrument contained nine items on sexual behavior and relationship. Participants responded to 1) the open-ended Questions 1 - 3 (e.g., With how many different partners have you had sex within the past 12 months?), 2) they responded on a 5-point Likert scale ranging from 1 - Strongly

disagree to 5 - Strongly agree for Questions 4 - 6 (e.g., Sex without love is OK), and 3) a 9-point Likert scale ranging from 1 – never to 5 – nearly every day for Questions 7 - 9 (e.g., How often do you have fantasies about having sex with someone you are not in a committed romantic relationship with?). Penke (2010) reported Cronbach's $\alpha = .83$ for all the items. For this present study, we used all the 9 items.

Demographic questionnaire. We then asked participants to complete demographic questions²⁰ concerning their gender, sexual orientation, level of education, age, country of origin, religion, parental status, marital and relationship status and years, and monthly income.²¹

7. Cultural adaptation and translation of our measures

When necessary (i.e., when requested by the local site), we adapted the materials to participants' cultures to provide a better test of the relevant phenomenon (for all requested changes, see our OSF page at <https://osf.io/52mrt/>). After the cultural adaptation, we translated the questionnaire into a country's lingua franca and/or one widely spoken language of contributing labs (again, if requested by the local site). All translators were bilingual and self-reported to have at least C1 competency level (Council of Europe, 2020) in English and in the target translation language. While translating the original materials, the translators identified wordings that had no meaning or were unfamiliar in their locality. We adopted the Adetula et al. (2021) 3-step double-forward translation procedure, which required four translators per language:

²⁰ In the Teenage Pregnancy Study, the original authors collected information on the participants' tribe. However, we removed the tribe question because it is too close to asking about race. The lead researchers are in France, and researchers are prohibited from asking about race in France.

²¹ We didn't report participants' income because the participants' reports of their income after taxes proved too messy to use.

1. **Forward translation:** two translators (identified as Translator A and Translator B) translated the measures from English into the targeted local language and produced Version A and Version B translated text, respectively.
2. **Translated text adjudication:** a different and third translator reviewed Version A and Version B, translated text statement-by-statement, and adjudicated and selected translated texts from both versions that were contextually appropriate, culturally meaningful, and preserved the original meaning and intent. The adjudicator then produced Version C.
3. **Proofreading:** The fourth and last translators proofread the final Version C. All translations are accessible at <https://osf.io/h9673/>.

8. Data analysis plan

Main analyses. We included data from all sites in the analysis, even if they collected data from less than 80 participants. Broadly, we aimed to 1) establish if each original effect replicated *across* sites while allowing for by-site variability across sites. Then, we 2) examined differences for each effect between participants from different continents, with participants in Africa as the comparison group (e.g., Africa versus Europe comparison and Africa versus North America comparison). Finally, we examined whether there was heterogeneity across our entire sample and within participants from each of the three continents.

We used the largest random effects structure justified by the design (Barr, Levy, Scheepers, & Tily, 2013) except in cases where the most maximal model was not identifiable. This meant that, in contrast to the original analyses, we included by-site random intercepts and any by-site random slopes justified by the design (see Table 9 above). In cases of non-identifiability, we trimmed the top-level by-site random interaction terms until identifiability was reached.

For all three studies, the nature of the original samples differed slightly from ours. Our main replication analyses were, therefore, subgroup analyses. For the original teenage pregnancy study our main analyses were subgroup analyses for parents aged 35 and older across the three continents²². For the condom use negotiation study, our main analyses consisted of subgroup analyses with heterosexual male and female college students between 18 - 20 years old across the three continents. For the original extramarital sexual relationship study, our main analyses consisted of subgroup analyses for heterosexual women across the three continents.

The full statistical models and analysis script are accessible at <https://osf.io/y8kb6>. We preregistered this current study. The preregistration document is available at <https://osf.io/pgvwk>.

²² However, very few participants from North America met these criteria. We, therefore, decided to exclude participants from this region for this study.

Table 13 Replication study statistical analyses deviations

Study	Original statistical analysis	Replication statistical analysis	Deviation justification
Teenage pregnancy	Used frequencies and percentages to describe the proportions of agreements to the statements of attitudes towards teenage pregnant girls	Used a generalized linear mixed effects model to estimate 1) the fixed intercept, 2) fixed effect for North American versus Africa, and 3) fixed effect for European versus Africa on attitudes towards teenage pregnancy	Mixed effect models are more appropriate to estimate the participants' attitudes within and between the regions because they account for by-site variance.
Condom use negotiation	Used one-way MANOVA of gender as a fixed factor on 5 variables of condom use as dependent variables, and a series of mean differences in responses between males' and females' on these factors independently.	Used a linear mixed effects model to estimate both the 1) the main effect and the 2) by-site variance for all of the following: men's versus women's a) attitude to condom use, b) control over condom use, and c) self-reported condom-protected sexual behaviors across Africans, Europeans, and North Americans.	Mixed effects models are appropriate because we treated the dependent variables individually, not as a composite dependent variable, and because they account for by-site variance.
Extramarital sexual relationship	Used analyses of variance (ANOVA) to test the main and interaction effects of emotion involvement, regret expressed, and commitment to marriage conditions on participants' willingness to forgive.	Used a linear mixed effects model to estimate the main effects of emotional involvement and regret expressed on willingness to forgive across all region and between Africa and Europe, and Africa and North America	Mixed effects models are a more appropriate because they account for by-site variance

Note. The statistical analyses and study design (for Extramarital Sex Study only) deviations from the original studies and justification for the deviations.

Auxiliary analyses. Because we sampled a broader group of participants than the original studies, we conducted additional analyses to examine our focal effects with our full group of participants, and for Africa versus Europe and Africa versus North America. We report these results in Appendixes VI, VII, and VIII to focus entirely on the subgroups in the results section which were the focal effects of this current replication.

9. Results

To ensure convergence, we checked which optimizers yielded convergence using `lme4::allFit`. For each study, we first present the results of the main subgroup analyses across the regions. We then provided our inferences regarding whether effects were replicated for these different comparisons. For these auxiliary analyses, we also compare between European versus African participants, then North American versus African participants

9.1 Teenage pregnancy study results

Only 24 of our participants from North America were parents 35 and above, so analyses didn't include this subgroup from North America. Without a specific hypothesis to test, we estimated the agreement proportions across all regions for individual 13 items and their CIs and tested whether the CIs overlap between the original and the replication samples. We present our findings below in 2 tables: Table 14, the estimated probability for the focal effect estimates for 13 items on attitudes towards teenage pregnancy, and their predicted probability CIs for parents aged 35 and above, and the original calculated CIs used for replication test if those overlap with replication predicted probability CIs. In Table 15, we present the fixed effect estimated probability and their predicted probability CIs comparing Africa versus Europe on these items, as well as the predicted probability for Europe and Africa and the *sd* for sites for the overall subgroup, Europe, and Africa.

9.1.1 Agreement proportions for parents aged 35 and above

The participants' attitudes towards teenage girls were supportive and positive. Participants disagreed with negative notions and actions (e.g., the pregnant teenage girl does not have a future or her parents have failed), agreed with positive ones (such as supporting the girl), and were indifferent about forcing the man/boy responsible for the pregnancy to take responsibility or whether the pregnancy situation devastates the parents (see Table 14).

9.1.2 Replication test of teenage pregnancy

We tested whether, across regions (Europe and Africa) and among parents 35 and above, the confidence intervals around the predicted probability of agreement overlapped with the confidence intervals from the original study. The replication test showed that only the original and replication CIs for Item 10 overlapped, while CIs for other Items did not overlap (see Table 14). Hence, we replicated the findings that most parents disagreed that “parents should abort their teenage daughter’s pregnancy.”

Table 14 Teenage pregnancy replication results for parents aged 35 and above, and the replication decision

Item/Statement of test	Estimated probability for $b_{intercept}$	95%CI predicted probability for $(b_{intercept})$ Lower, Upper	Original study 95%CI for agreement Lower, Upper	Replication decision	$sd(b_{intercept})$
<i>Item 1: A pregnant teenage girl is a deviant</i>	0.149	0.075, 0.275	0.904, 0.960	Did not overlap, Not replicated	1.56
<i>Item 2: A pregnant teenage girl do not have future</i>	0.067	0.048, 0.093	0.797, 0.879	Did not overlap, Not replicated	0.00
<i>Item 3: Teenage pregnancy makes parents devastated</i>	0.550	0.381, 0.708	0.819, 0.897	Did not overlap, Not replicated	1.44
<i>Item 4: Teenage pregnancy means that parents have failed</i>	0.155	0.105, 0.223	0.815, 0.894	Did not overlap, Not replicated	0.76
<i>Item 5: Teenage pregnancy results in stigmatization</i>	0.719	0.653, 0.776	0.859, 0.928	Did not overlap, Not replicated	0.34
<i>Item 6: Teenage pregnancy is the result of western education</i>	0.132	0.085, 0.200	0.651, 0.753	Did not overlap, Not replicated	0.72
<i>Item 7: Parents should drive girl away from the family in anger</i>	0.017	0.006, 0.046	0.275, 0.379	Did not overlap, Not replicated	0.76
<i>Item 8: Parents should stop her schooling</i>	0.023	0.010, 0.051	0.442, 0.554	Did not overlap, Not replicated	0.74
<i>Item 9: Parents should give her out in early marriage</i>	0.053	0.030, 0.092	0.574, 0.682	Did not overlap, Not replicated	0.75
<i>Item 10: Parents should abort pregnancy</i>	0.121	0.079, 0.179	0.118, 0.199	Overlapped, Replicated	0.59
<i>Item 11: Parents should send her to a maternity group home for pregnant teenage girls for care</i>	0.137	0.080, 0.223	0.644, 0.747	Did not overlap, Not replicated	1.10

Increasing the generalizability of psychology by including Africans

Item 12: <i>Parents should support girl to have the baby</i>	0.839	0.772, 0.889	0.521, 0.631	Did not overlap, Not replicated	0.68
Item 13: <i>Parents should force the man/boy responsible for the pregnancy to accept responsibility of her care</i>	0.421	0.333, 0.515	0.801, 0.882	Did not overlap, Not replicated	0.63

Note. A summary table showing the estimated probability of the focal effect for the ($b_{intercept}$) for parents aged 35 and above, and their predicted probability CIs and standard deviations. For the replication decision, when we compared the replication CIs predicted probability and original CIs. Only Item 10 CIs for original and replication overlapped, hence were replicated.

9.1.3 Region comparisons for agreement proportions for parents aged 35 and above

To compare the regions' proportion agreements for these 13 items, we fitted dummy-coded indicator variables for the comparisons between Africa and Europe as fixed slopes in our model (see Table 9) using Africa as base for comparison. The estimated probability of 11 out of the 13 fixed effects for Africa versus Europe were significant (see Table 15). African parents' attitudes towards pregnant girls differed from those of Europeans on the following items: 1) a pregnant teenage girl a deviant, 2) a pregnant teenage girl do not have future, 3) teenage pregnancy makes parents devastated, 4) teenage pregnancy means that parents have failed, 5) teenage pregnancy results in stigmatization, 6) teenage pregnancy is the result of Western education, 7) parents should drive girl away from the family in anger, 8) parents should stop her schooling, 9) parents should give her away in early marriage, 10) parents should send her to a maternity group home for pregnant teenage girls for care, and 11) parents should force the man/boy responsible for the pregnancy to accept responsibility of her care. The attitudes of African parents did not differ from those of Europeans on the following items: 1) parents should abort pregnancy, and 2) Parents should support girl to have the baby.

Although African and European parents aged 35 and above shared similar opinions about supporting a teenage pregnant girl rather than aborting the baby, sending her away from home, or stopping her from school, African parents are more likely to see the pregnant teenager as a deviant and could face stigmatization, and that the pregnancy could jeopardize her future. African parents compared to European parents would be more devastated, see themselves as having failed as parents, blame Western education, and are more likely to give away their pregnant teens in early marriage, send her to a maternity home for care, or even force the boy or man responsible for the pregnancy to take responsibility for her care.

We ran all the focal effects, replication and region comparisons tests for the full sample that include participants outside the parents aged 35 years and above, and the North America sample. Hence, we had a much larger sample for the full sample analysis. The agreement probabilities were consistent with subgroup findings. Also, we only replicated Item 10 which most parents disagreed with “parents should abort their teenage daughter’s pregnancy.” Expectedly the region comparisons were inconsistent with the subgroup findings for some variables because, in addition to the full sample comparison for Africa versus Europe, we had Africa versus North America comparison (see full report in Appendix VI).

Table 15 Teenage pregnancy replication region comparisons for parents aged 35 and above, and heterogeneity test

Item of test	Africa v. Europe				Random effects		
	Estimate	95%CI Lower, Upper	<i>Predicted</i> % (<i>Europe</i>)	<i>Predicted</i> % (<i>Africa</i>)	<i>sd</i> _{Site(Overall)}	<i>sd</i> _{Site(Europe)}	<i>sd</i> _{Site(Africa)}
<i>Item 1: A pregnant teenage girl is a deviant</i>	10.73**	2.35, 49.04	0.051	0.365	1.56	2.31	0.77
<i>Item 2: A pregnant teenage girl do not have future</i>	4.58***	2.25, 9.32	0.033	0.134	0.00	0.00	0.00
<i>Item 3: Teenage pregnancy makes parents devastated</i>	16.83***	4.03, 70.37	0.229	0.834	1.44	2.09	0.48
<i>Item 4: Teenage pregnancy means that parents have failed</i>	2.68*	1.09, 6.56	0.101	0.231	0.76	0.96	0.06
<i>Item 5: Teenage pregnancy results in stigmatization</i>	2.14*	1.16, 3.94	0.636	0.789	0.37	0.41	0.00
<i>Item 6: Teenage pregnancy is the result of western education</i>	3.49**	1.44, 8.46	0.075	0.222	0.72	0.35	1.35
<i>Item 7: Parents should drive girl away from the family in anger</i>	11.38**	2.34, 55.34	0.005	0.055	0.76	0.70	0.79
<i>Item 8: Parents should stop her schooling</i>	7.59**	1.95, 29.52	0.008	0.061	0.74	1.17	0.02
<i>Item 9: Parents should give her out in early marriage</i>	7.20***	2.43, 21.38	0.021	0.132	0.75	1.18	0.00
<i>Item 10: Parents should abort pregnancy</i>	1.16	0.47, 2.91	0.113	0.129	0.59	0.36	0.00
<i>Item 11: Parents should send her to a maternity group home for pregnant teenage girls for care</i>	10.67***	3.18, 35.84	0.046	0.341	1.10	1.43	0.73

Increasing the generalizability of psychology by including Africans

Item 12: <i>Parents should support girl to have the baby</i>	0.91	0.38, 2.17	0.845	0.832	0.68	0.75	0.39
Item 13: <i>Parents should force the man/boy responsible for the pregnancy to accept responsibility of her care</i>	5.11***	2.43, 10.73	0.244	0.622	0.63	0.66	0.52

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

A summary table showing the estimates of the fixed effects for the subgroup between regions for Africa versus Europe for the 13 items on attitude towards teenage pregnancy and their %CIs, the predicted probability per region, and by-site standard deviations. Also, the by-site standard deviations for Africa and Europe subgroups for the within region heterogeneity test.

9.2 Condom use negotiation study results

We hypothesized that male participants had 1) more favorable attitudes towards condom use, 2) greater feelings of control over condom use, and 3) higher self-reported condom-protected sexual behaviors, than their female counterparts. We present our findings in Table 16 for the focal effect estimates of gender on attitudes towards condom use, control over condom use, and condom-protected sexual behavior, and its fixed effect estimates comparing Africa versus Europe and Africa versus North America, and their CIs for heterosexual male and female college students aged 18 to 20, and lastly, the replication test and decision.

9.2.1 Main effect of gender on the outcome variables

When we used our model (see Table 9) to test the focal effects, namely the difference in each variable across heterosexual male and female students aged 18 to 20 years old. Although both male and female adolescents showed positive attitudes (Figure 5) and considerable control over condom use (Figure 6), there was no significant difference between male and female students reported attitude towards condom use ($b_{gender} = .01$; $t_{gender}(34.64) = .06$, $p = .954$, $sd(b_{gender}) = .48$), and control over condom use ($b_{gender} = -.15$; $t_{gender}(180.40) = -1.59$, $p = .114$, $sd(b_{gender}) = .12$). However, as shown in Figure 7, male college students were more likely to report engaging in condom-protected behavior more than their female counterparts ($b_{gender} = -.48$; $t_{gender}(35.44) = -4.05$, $p = .0003$, $sd(b_{gender}) = .25$). This implied that male and female adolescent students did not differ in their attitudes toward and perceived control over condom use, but did in their self-reported condom-protected sexual behavior.

9.2.2 Replication test for condom use negotiation focal effects

To test if we replicated the fact that adolescent male college students had 1) more positive attitudes towards condom use, 2) greater perceived control over condom use, and

higher self-reported condom-protected sexual behavior than their female counterparts, we tested whether the main effects of gender on these outcome variables' p -values were larger or smaller than the preset p -value of .0167 (see Table 16). We determined that only the main effect of gender on condom-protected sexual behavior was replicated, while the effects of gender on attitudes and perceived control were not.

9.2.3 Gender by region comparison interaction

To test whether the focal effects differ by region, we fitted dummy-coded indicator variables for the comparisons between Africa and either Europe or North America as fixed slopes in our model (see Table 9) using Africa as base for comparison and allowed these to interact with gender. The reported condom-protected sexual behavior mean difference ($MD = 0.91$) between African male students' ($M = 3.53$; $sd = 1.51$) and female students' ($M = 2.62$; $sd = 1.21$) was larger compared to European male students' ($M = 3.36$; $sd = 1.55$) and female students' ($M = 3.24$; $sd = 1.23$) mean difference ($MD = 0.12$, $b_{gender*afrika_v_eu} = 0.87$; $t_{gender*afrika_v_eu}(52.36) = 2.61$, $p = .012$, Table 16). While all male students reported more condom-protected sexual behavior than female students, this was especially true for African male students, as compared to European male students. Other gender differences by region tests were non-significant (see Table 16).

The full sample analysis of condom use negotiation showed that the focal effects, replication tests, and region comparisons results were not consistent with our subgroup findings. Contrary to our subgroup results in which only the self-reported condom protected behavior was significant, we found that the main effects of gender across the entire participants and its interaction effects for the region comparisons were significant for all the variables except for attitude towards condom use for Africa versus North America (see Appendix VII).

Table 16 Condom use negotiation results for the subgroup analyses, and the replication test

Outcome variable	Fixed effects		95%CI		SE	t-value	df	p-value	Random effects				Replication test		
	Parameter	Original estimate	Replication estimate	Lower Bound					Upper Bound	sd _{Site(O} verall)	sd _{Site(E} urope)	sd _{Site(Af} rica)	sd _{Site(Nort} h America)	Preset <i>a</i>	Replication decision
Attitude	$(b_{gender})_{18-20}$	-0.23	0.01	-0.29	0.31	0.15	0.06	34.64	.954	0.48	0.63	0.41	0.29	0.0167	Non replication
	$(b_{gender*africa_v_eu})_{18-20}$	-	0.04	-0.76	0.84	0.39	0.09	50.55	.927	0.48	-	-	-	-	
	$(b_{gender*africa_v_us})_{18-20}$	-	0.11	-0.64	0.87	0.37	0.30	34.21	.759	0.48	-	-	-	-	
Control	$(b_{gender})_{18-20}$	-0.34	-0.15	-0.35	0.04	0.09	-1.59	180.40	.114	0.12	0.08	0.12	0.12	0.0167	Non replication
	$(b_{gender*africa_v_eu})_{18-20}$	-	0.34	-0.21	0.88	0.28	1.22	270.72	.223	0.12	-	-	-	-	
	$(b_{gender*africa_v_us})_{18-20}$	-	0.33	-0.15	0.82	0.25	1.35	167.68	.180	0.12	-	-	-	-	

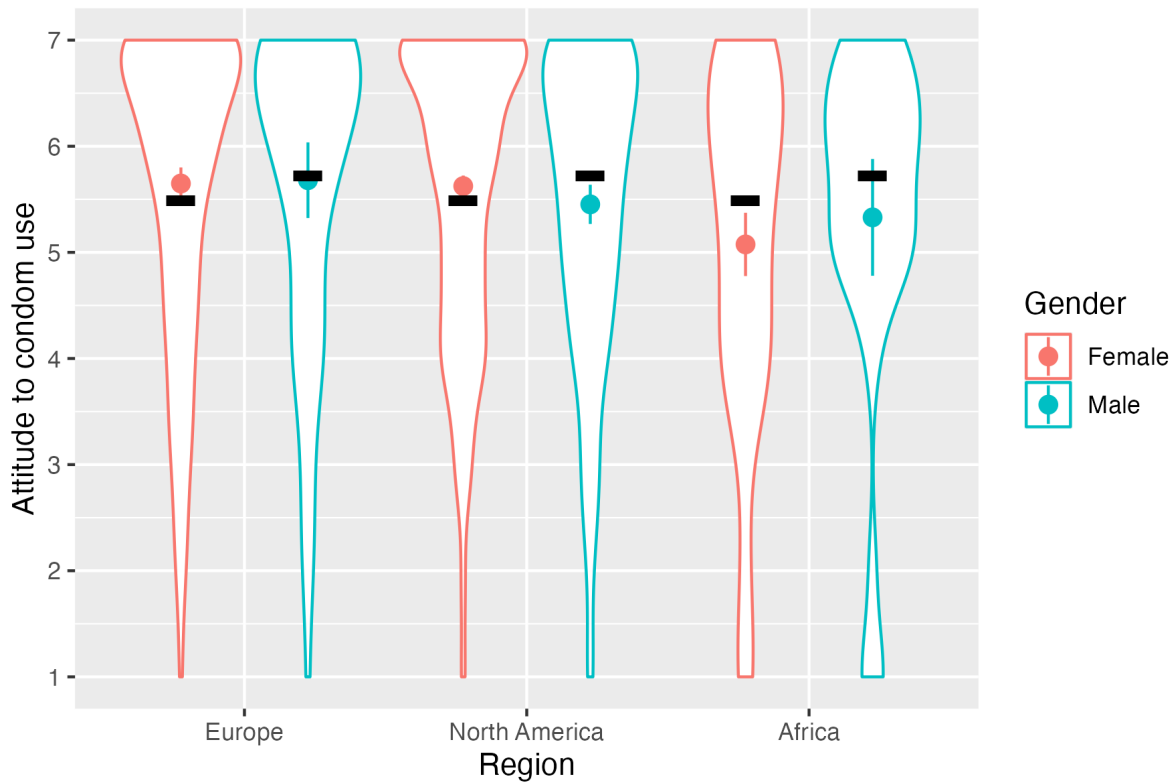
Increasing the generalizability of psychology by including Africans

Protected sexual behavior	$(b_{gender})_{18-20}$	-0.74	-0.48	-0.72	-0.24	0.12	-4.05	35.44	.0003	0.25	0.44	0.32	0.13	0.0167	Replication
	$(b_{gender*africa_v_eu})_{18-20}$	-	0.87	0.20	1.53	0.33	2.61	52.36	.012	0.25	-	-	-	-	
	$(b_{gender*africa_v_us})_{18-20}$	-	0.55	-0.06	1.15	0.30	1.82	36.68	.077	0.25	-	-	-	-	

Note. The table provides a summary of the original fixed effect estimate derived from the male and female mean differences for the three outcome variables. The replication fixed effects estimates of gender on attitudes towards to condom use (“attitude”), perceived control over condom use (“control”), and self-reported condom-protected sexual behavior (“protected sexual behavior”) for heterosexual male and female college students between 18 - 20 years, and for Africa versus Europe, and for Africa versus North America, and as well as by-site variability for these effects, and their CIs.

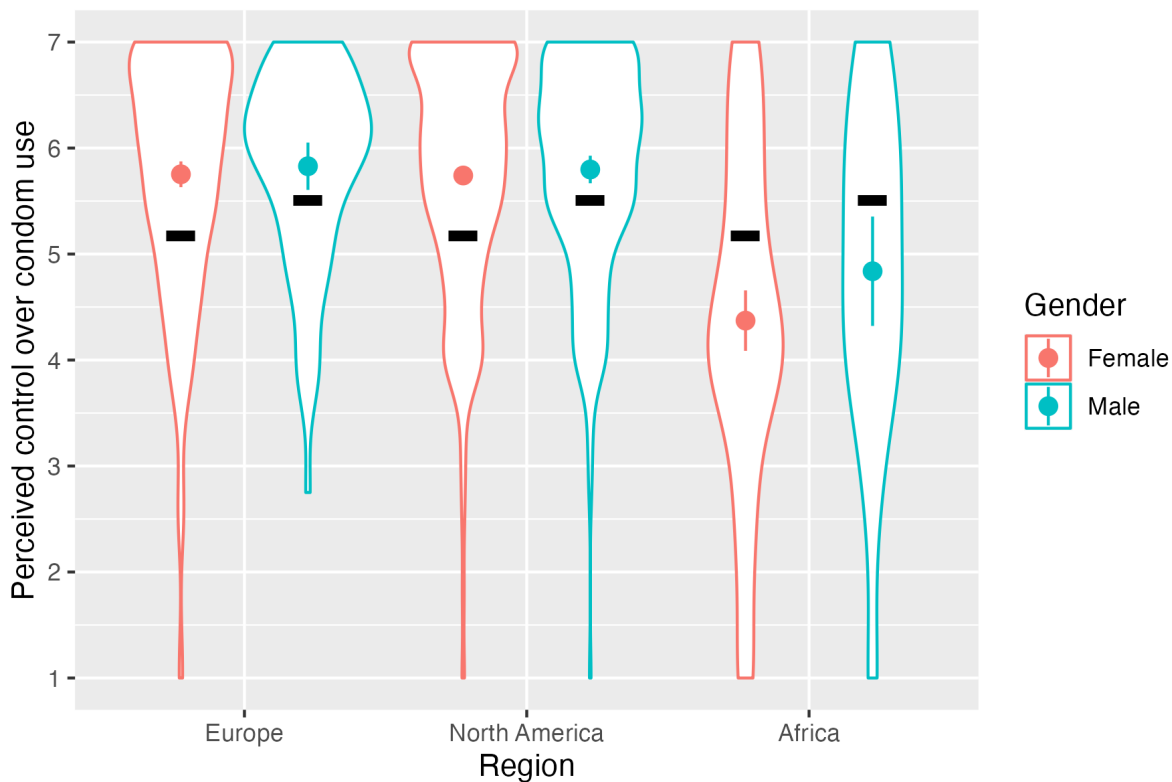
Given the obtained p -values for the main effects for gender on attitude towards condom use ($b = .01$, p -value = .954), control over condom use ($b = -.15$, p -value = .114), and self-reported condom-protected sexual behavior ($b = -.48$, p -value = .0003), we compared the replication p -values to our preset alpha (0.0167), and found a replication for the effect of gender on condom-protected sexual behavior only.

Figure 5 Attitude to condom use by gender per region for adolescent students



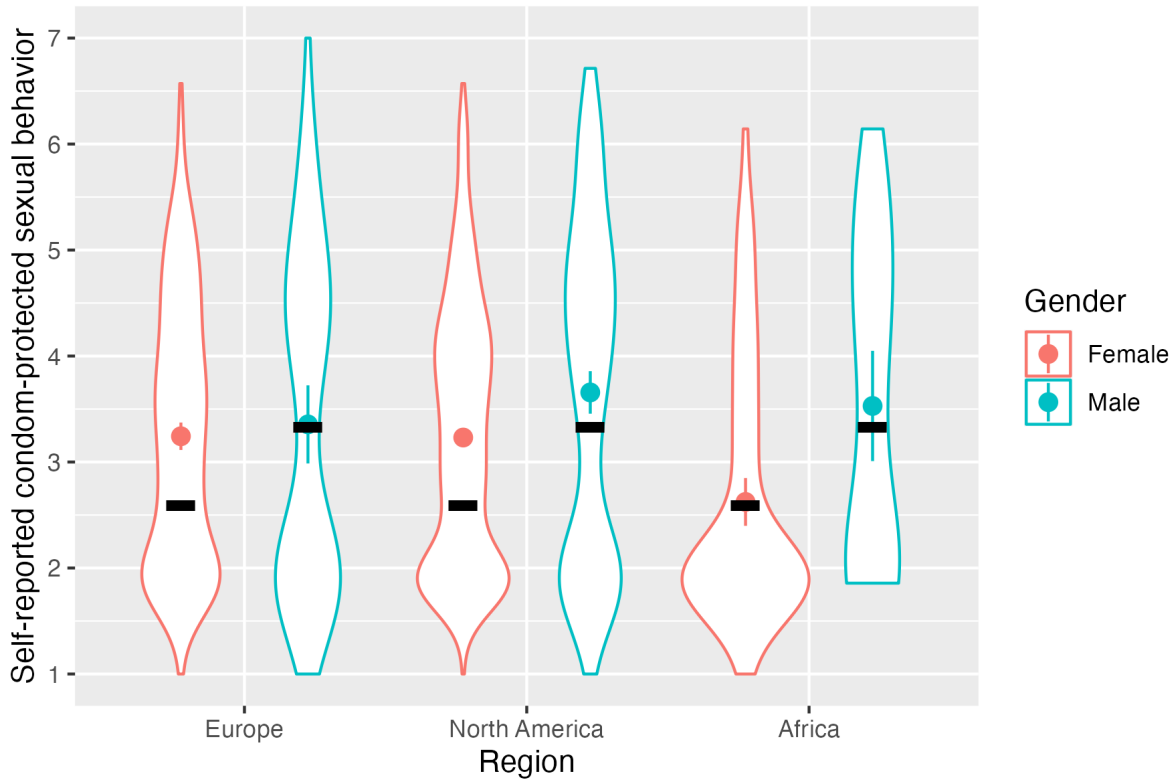
Note. Across all the regions of Europe, North America, and Africa, heterosexual male students aged 18 - 20 years, the average scores on attitude towards condom use were marginal between male and female, even when we compare African students to their male and female European and North American counterparts. In each violin are the original study male and female students' attitude towards condom use average scores (black crossbars). The original study attitude towards condom use means for African adolescent both male and female college students are larger than our replication mean scores for Africans but marginally similar to our replication Europeans and North Americans male and female students average scores.

Figure 6 Perceived control over condom use by gender per region for adolescent students



Note. The average scores on perceived control over condom use were marginal between heterosexual male (blue dots) and female (red dots) students aged 18 - 20 years for both Europeans and North Americans, but considerable between male and female African students. Also, African male and female students mean scores are lesser compared to their European and North American counterparts. In each violin are the original study of male and female students' perceived control over condom use average scores (black crossbars). The original mean scores are larger than our replication mean scores for Africans but lesser than Europeans and North Americans male and female average scores.

Figure 7 Self-reported condom-protected sex behavior by gender per region for adolescent students



Note. Across all the regions of Europe, North America, and Africa, heterosexual male students aged 18 - 20 years average scores on self-reported condom-protected sexual behavior (blue dots) were higher than for their female counterparts (red dots). Also, in each violin plot for all regions are the original study male and female students' self-reported condom-protected sexual behavior average scores (black crossbars). The original study self-reported condom-protected sexual behavior means for adolescent male college students are lower than our replication mean scores for male and female students across all regions except for the European male and African female student categories.

9.3 Extramarital sexual relationship study results

Although the full design for the Vera Cruz replication is a 2x2x2 factorial design of low versus high degrees for emotional involvement, regret expressed, and commitment to marriage, however, we tested only two hypotheses for the focal effects, and reported the full results in Appendix VII. We tested whether women were more willing to forgive 1) a husband with a low degree of emotional involvement in an extramarital affair compared to a husband with a high degree emotional involvement in an extramarital affair, and 2) when a husband expressed high regret compared to when a husband expressed low regret for engaging in an extramarital sexual relationship. We present the extramarital sexual relationship study results in Table 17 for the focal main effect estimates of emotional involvement and regret expressed on willingness to forgive for heterosexual women as well as the replication tests.

9.3.1 Main effect of emotional involvement and regret expressed on forgiveness

In the original study, Mozambican women are more willing to forgive their partners when they express low emotional involvement in the extramarital sexual relationship and profound regret or high commitment to their relationship/marriage and vice versa. In this current replication study, we applied our linear mixed effects model (see Table 9) to examine this hypothesis among our heterosexual women respondents. Although they reported low willingness to forgive a cheating partner regardless whether the regret expressed or emotional involvement was low or high ($M_{\text{Africa}} = 2.54$, $sd_{\text{Africa}} = 3.00$; $M_{\text{Europe}} = 0.84$, $sd_{\text{Europe}} = 1.60$; $M_{\text{North America}} = 0.89$, $sd_{\text{North America}} = 1.54$), heterosexual women are more likely to forgive a cheating husband 1) that expressed high regret compared to one that expressed low regret ($b_{\text{regret}} = 1.17$, $p < .05$) and 2) less emotionally involved than one that is highly emotionally involved in an extramarital affair ($b_{\text{emotion}} = -.21$, $p < .05$, see Table 20).

Table 17 Extramarital sexual relationship focal effect results for heterosexual women only

Fixed effects								Random effects								Replication test	
Parameter	Original estimate	Replication estimate	95% CI		<i>t</i> -value	<i>df</i>	<i>p</i> -value	Overall		Europe		Africa		North America		Preset <i>a</i>	Replication decision
			Lower Bound	Upper Bound				<i>sd</i> _{ID}	<i>sd</i> _{Site}	<i>sd</i> _{ID}	<i>sd</i> _{Site}	<i>sd</i> _{ID}	<i>sd</i> _{Site}	<i>sd</i> _{ID}	<i>sd</i> _{Site}		
<i>b_{emotion}</i>	-1.97	-.21	-0.26	-0.15	-8.58	28.38	.00000000	0.67	0.09	0.76	0.13	0.65	0.04	0.52	0.02	.0005	Replication
<i>b_{regret}</i>	4.21	1.17	0.98	1.36	12.46	33.71	.00000000	1.41	0.55	1.24	0.21	2.13	0.93	1.05	0.13	.0005	Replication
<i>b_{emotion*africa_v_eu}</i>	-	0.20	-0.32	-0.08	-3.47	33.47	.001	0.67	0.09	-	-	-	-	-	-	-	-
<i>b_{regret*africa_v_eu}</i>	-	-1.25	-1.69	-0.82	-5.82	36.01	.000001	1.41	0.55	-	-	-	-	-	-	-	-
<i>b_{emotion*africa_v_us}</i>	-	-0.14	-0.27	-0.01	-2.22	30.26	.034	0.67	0.09	-	-	-	-	-	-	-	-
<i>b_{regret*africa_v_us}</i>	-	-1.34	-1.83	-0.84	-5.45	34.65	.000004	1.41	0.55	-	-	-	-	-	-	-	-

Note. Dependent variable - Willingness to forgive

Original and replication main effect estimates of emotional involvement and regret expressed on willingness to forgive for heterosexual women participants only, as well as their interactions with Africa_vs_Europe and Africa_vs_North America. We also present the by-person and by-site standard deviations for these focal effects for the overall participants and for participants in each region. Lastly, possible replication interpretations of our results given *p*-values we obtain for the main effects for emotional involvement ($b_{emotion} = -.21, p\text{-value} = .000000002$) and regret expressed ($b_{regret} = 1.17, p\text{-value} = .000000000000003$) for heterosexual women compared to the original study preset alpha (.0005).

9.3.2 Replication test for extramarital sexual relationship focal effects

To determine whether we were able to replicate Vera Cruz (2018), we focused only on p -value because we reduced the degrees/levels for both emotional involvement and regret expressed which means there are no bases for confidence interval comparisons. Our replication test showed these estimates' p -values were smaller than the original study preset p -value (see Table 17). Hence, we replicate these effects. However, although the replication results are in the same direction as the original study, the effect sizes are much less strong.

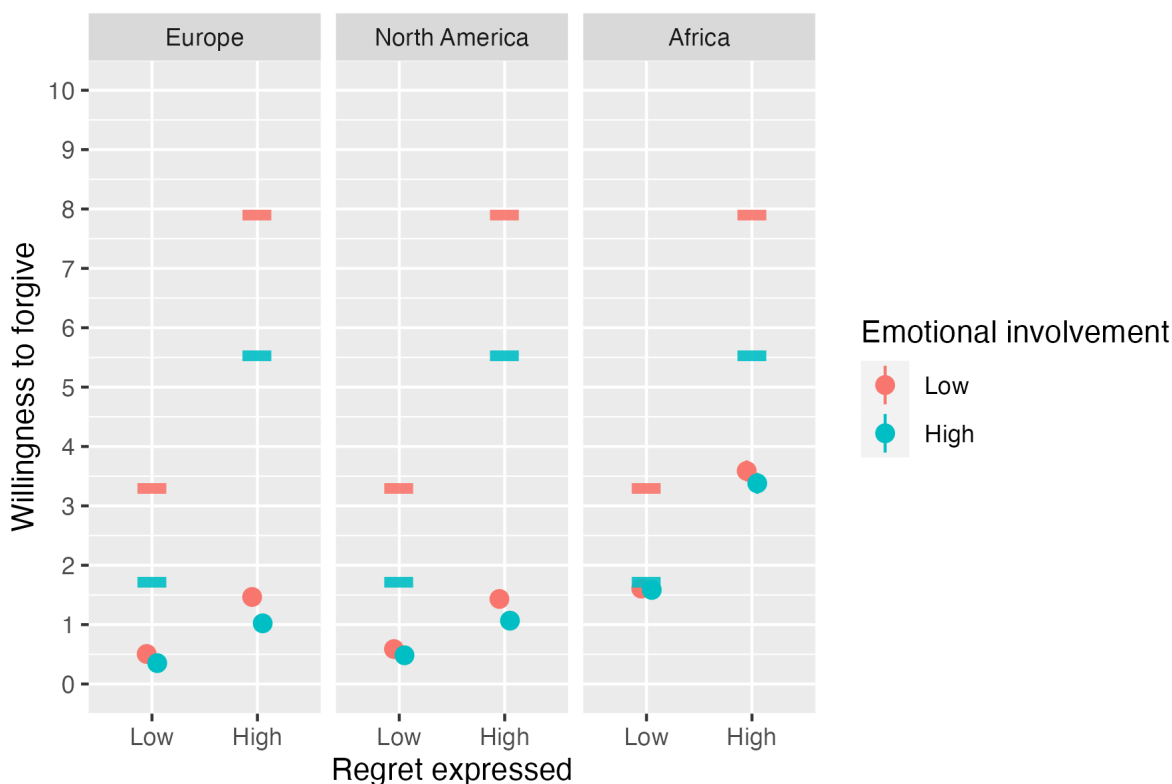
9.3.3 Region comparison

To test whether the focal effects differ by region, we fitted dummy-coded indicator variables for the comparisons between Africa and either Europe or North America as fixed slopes in our model (see Table 9) using Africa as base for comparison to test the interaction effects of emotional involvement and regret expressed and region. The impact of regret expressed is much larger among African women compared to European women ($b_{regret*afrika_v_eu} = -1.25, p < .05$) and North American women ($b_{regret*afrika_v_us} = -1.34, p < .05$). When the regret expressed was low, African women ($M = 1.59, sd = 2.44$) were more willing to forgive a cheating husband than their European ($M = 0.43, sd = 1.11$) and North American counterparts ($M = 0.54, sd = 1.23$). Similarly, when the regret expressed was high, African women ($M = 3.48, sd = 3.20$) were more willing to forgive a cheating husband compared to European ($b_{regret*afrika_v_us} = -1.34, p < .05; M = 1.24, sd = 1.89$) or North American women ($M = 1.25, sd = 1.73$).

In Europe ($b_{regret*afrika_v_eu} = .20, p < .05$) and North America ($b_{regret*afrika_v_us} = -.14, p < .05$), women were slightly less willing to forgive the affair when emotional involvement was high, but involvement doesn't make as much of a difference in Africa. When the emotional

involvement was low, African women ($M = 2.59, sd = 3.02$) were equally as willing to forgive a cheating husband when the emotional involvement was high ($M = 2.48, sd = 2.98$). Whereas when the emotional involvement was low, European ($M = 0.99, sd = 1.74$) and North American women ($M = 1.01, sd = 1.61$) were marginally more willing to forgive a cheating husband when the emotional involvement was high ($M_{\text{Europe}} = 0.69, sd_{\text{Europe}} = 1.43; M_{\text{North America}} = 0.78, sd_{\text{North America}} = 1.48$; see Figure 8). The extramarital sexual relationship focal effects, replication tests, and region comparisons results were consistent with our entire participant results (see Appendix VII).

Figure 8 Willingness to forgive mean ratings as an operation of regret expressed and emotional involvement and for region comparison



Note. Willingness to forgive average ratings on the y-axis. Three panels correspond to the regions: Europe, North America, and Africa. For each panel we have regret expressed on the x-axis for low and high regret expressed conditions. In each panel, we have red and blue cross bars and dots indicating the combination of low versus high conditions for regret expressed and emotional involvement for the original study and replication respectively. The red cross bars/dots indicate low emotional involvement and the blue cross bars/dots high emotional involvement conditions. In the replication study, across the conditions and regions, when the regret expressed is high and the emotional involvement is low, participants are more likely to

forgive a cheating husband compared to when the regret expressed is low and the emotional involvement is high. Our replication average forgiveness ratings are much lower than reported in the original study²³.

9.4 Auxiliary analyses results

9.4.1 Heterogeneity across entire participants and within regions

We anticipated variations across participants and within regions for the focal effects due to cultural differences (e.g., site language used for study), participants' recruitment procedure (e.g., online versus in-person settings), and countries'/regions' disproportionate sample size (e.g., 9 participants were sampled from Burkina Faso versus 1657 from the United States of America). To examine the heterogeneity across the regions, we drew on the focal effect estimates and their standard deviations for the subgroup analyses. However, to examine heterogeneity within regions, we subset the subgroup datasets into regions (for Africa, for Europe and for North America) and applied our models (see Table 9). We estimated the by-site and by-participant (for extramarital sexual relationship study only) standard deviations for each region separately. Aside from the main effects of gender on attitudes towards condom use, we observed that the by-site estimates' standard deviations for the intercepts of the 13 items for attitudes towards teenage pregnancy, for the main effects of gender on control over condom use, condom-protected sexual behavior, and for main effects of emotional involvement and regret expressed (except for the North America) on willingness to forgive were smaller than the focus effect estimates for these variable outcomes. Hence, we observed a relatively low amount of region-specific variability between sites (see Tables 15, 16, and 17).

Also, we examined the by-participant variability among participants for the extramarital sexual relationship study, since that study took multiple measurements per participant. We

²³ We extracted the original forgiveness means from the graph presented in Vera Cruz's (2018, Figure 1) using WebPlotDigitizer (Rohatgi, 2024; <https://automeris.io/WebPlotDigitizer.html>).

observed that by-participant slopes for emotional involvement and for regret expressed were larger than the main effect estimates for the subgroup (heterosexual women), for Africans, for Europeans, and for North Americans. Hence, we found a relatively large amount of variability across the participants for the subgroup as well as within regions (see Table 17). All these focal effect heterogeneity results for these subgroups were consistent for the full participants (see Tables 31, 32, and 33 in Appendixes VI, VII, and VIII, respectively).

10. Discussion, Recommendation, and Conclusion

10.1 Discussion

Contrary to the big team science approach of generalizing to Africa, we set out to generalize three claims originally discovered in Africa to Africa, Europe, and North America. Guided by our Global South experts' opinions, we focused on social applied issues and adopted a study selection approach that allowed African collaborators to provide vital input in selecting studies of interest. Our collaborators' assessment of the five shortlisted studies after rounds of study nomination, search, and evaluation showed that, compared to European and North American assessors, our African assessors were interested in reproductive health, sexual behavior and relationships studies (see Chapter 3a for the full report). Specifically, this current study replicated studies on attitudes toward teenage pregnancy (13 focal effects, originally studied among Nigerian parents 35 years and above), condom use negotiation (3 focal effects, originally studied among adolescent Ghanaian college students), and extramarital sexual relationships (2 focal effects, originally replicated among Mozambican women). Although the original sample characteristics of these three studies are distinct, the sexual behavior theme made for a good-fit bundled study, as we could sample a general adult (18 years and above) population and examine the studies' focal effects with subgroups similar to the original samples and also for region comparisons using Africa as the base.

Taken together, we replicated 4 out of the 18 focal effects, although some of the effects observed were much weaker than in the original studies. This investigation revealed that only these 4 African claims were clearly generalizable across Africans, Europeans and North Americans. Our region comparisons showed that the focal effect sizes differed between Africa on the one hand and Europe and North America on the other for 11 of the 13 attitudinal statements on teen pregnancy, the effect of gender on attitude towards condom use, perceived control over condom use, and self-reported condom-protected sexual behavior, and women's willingness to forgive hypothetical sexual infidelity. Our heterogeneity checks of the locality where these investigations were conducted showed that these findings do not vary between the sites across the continents and even for sites within the same continent.

For the teenage pregnancy replication, we observed that only Item 10 (parents should abort their teenage daughter's pregnancy) out of the 13 attitudinal statements was replicated and generalized to Africans and Europeans. For the other 12 attitudinal statements not replicated, the original study reported higher agreement proportions than what we observed in our replication and a lower agreement proportions for the Item 12 (parents should support girls to have the baby). Our African and European parents aged 35 years and above reported more positive attitudes towards teenage pregnant girls than those of the original sample (Nigerian parents aged 35 years and above, from Akpabuyo Local Government Area in Cross River state). Consistently across Africa and Europe, our participants showed positive attitudes toward pregnant teenage girls and that parents should support their teenage pregnant daughters (see similar findings from Bałanda-Bałdyga et al., 2020). Besides, they agreed that pregnant teenagers do face stigmatization. Africans' agreement proportions were larger than Europeans' on all the variables except for "parents should abort their teenage daughter's pregnancy" (see Table 15). These findings indicate a need for some interventions to manage a teenage pregnancy situation that could devastate parents, an item "Teenage pregnancy makes parents

devastated” that generalizes better to Africans than Europeans, and that has an adverse impact on teenage mothers’ future. Interventions should be directed at minimizing stigmatization and social rejection, as well as supporting pregnant teenage girls and their parents (see also Ajayi et al. 2023; Cook & Cameron, 2015; Webb et al., 2023)

For the condom use negotiation replication, we were only able to replicate and generalize the effect of gender on self-reported condom-protected sexual behavior among male and female adolescent college students aged 18 to 20 years old in Africa, Europe, and North America. This finding generalizes more to our African subgroup than their European and North American counterparts. Compared to the original study, while the original average ratings on self-reported condom-protected sexual behavior for Ghanaian adolescent male college students 14 to 20 years old were slightly lower than our male students aged 18 to 20 years average ratings (even similar to our European male students), their female counterparts’ average ratings on this variable were considerably lower than our European and North American female categories but similar to our African female group. Nevertheless, our participants reported a moderate level of self-reported condom-protected sexual behavior in the past three months, which was consistent with findings on unsafe sexual behavior studies investigated among 15 to 24 years old male and female Malawians (Jimu et al., 2023) and 14 to 25 years old females in the United States (Nathan et al. 2023). Although our participants reported a more favorable attitude towards condom use and greater control to use condom, the gender difference in these variables was not different across regions. Our male students reported a higher average on self-reported condom-protected sexual behavior than their female counterparts that did not vary much across the three regions. Compared to their European and North American counterparts, our African women reported the lowest level of condom use. This may particularly be true for young African women in “less sexually liberal or permissive” African cultures who may find

it embarrassing to be proactive in accessing contraceptives or to negotiate condom-protected sex (see also findings from Agu et al., 2022; Flanagan et al., 2021).

For the extramarital sexual relationship study, we replicated whether husbands' emotional involvement in extramarital affairs and regret expressed by husbands impact participants' willingness to forgive the husbands. Across the board, participants reported very low willingness to forgive a cheating partner with an average rating of 1.24 (out of a total of 10 points). African heterosexual women are more willing to forgive a cheating husband than their European and North American counterparts. Mozambican women's (the original sample) forgiveness ratings were much higher than our heterosexual women's forgiveness average ratings, these findings generalize better to our African subgroup than their European and North American counterparts. Although the heterogeneity test revealed that our participants' forgiveness ratings varied across and within the continents, the effects of regret expressed and emotional involvement on forgiveness were consistent across the region, and when we compared Africans to Europeans, and to North Americans. Consistent with the original study, we found that women were willing to forgive when emotional involvement is low (versus high), and when regret expressed by a husband is high (versus low). This finding is consistent with Meskó & Láng (2013) that women expressed more distress, greater threat to their relationships, and were more willing to forgive an emotional infidelity with a prostitute than with an ex-girlfriend.

Beyond generalizability and theoretical advancement, this current replication study can benefit other research discourse. We discuss the low replication rate in this current study, how this current study facilitates capacity building, the difficulties conducting this study, and how the ManyLabs Africa replication compares to others.

Low replication rate. Only 4 out of the 18 African claims replicated and generalized to other Africans, Europeans, and North Americans. This study cannot distinguish between

different explanations for the low replication rate. Instead we discuss broadly the high non-replication rate in light of cultural variations and questionable research practices. One issue that could explain the low replication and generalization is cultural differences that exist between an African population and Europe/North America. For instance, societies that are more individualistic or collectivistic differ from each other in social norms and values (Hofstede et al. 2010), which are related to how sexual relationships are addressed in a specific population. Individualism has been correlated with sexual behavior and has even shown a positive correlation to sexual frequency (Ubillos et al. 2000). These cultural differences may be especially profound in the sexual behaviors (Ubillos, et al., 2000) investigated in this study because African cultures are less sexually liberal or permissive compared to the North American and European cultures (see also Ahlberg, 1994).

Another issue especially for low replication rate could be the quality of the original studies we replicated. Although little is known about the prevalence of questionable research practices or misconduct in African research (Kombe, et al, 2014), questionable research practices in original studies have been put forward as a possible explanation for replication failures of Western psychology studies (Munafò et al., 2017). These practices may also affect African psychology. These practices allowed researchers the flexibility of best-fit methods and practices to test an hypothesis leading to, for instance *p*-hacking, a use of various strategies to increase the chances of finding statistically significant results (Ioannidis, 2005) or Hypothesizing after results are known (HARKing), presenting an unpredicted finding as a priori prediction (Kerr, 1998), practices linked to replication failures. The teenage pregnancy study that accounted for 12 out of the 13 claims that failed our replication. For the teenage pregnancy study, there were no clearly stated hypotheses and how the participants' responses were quantified was poorly documented which limited the extent to which we can work with the study materials. We had to focus on the simple agreement proportions. Also, during the

study selection, we dropped our initial extramarital affairs study (Ojedokun, 2015) because we judged the measure used to be poorly developed. We do not bring up these cases to imply that questionable research practices are widespread in psychology. In fact, the Vera Cruz (2018) extramarital sexual relationship replicated and generalized to European and North America, suggesting that the original study was conducted reproducibly. However, African research can improve rigor and transparency by adopting credibility-enhancing practices. We discuss some of these and their application to African psychology in Chapter 5.

Building capacity. As a follow-up on the CREP study, this current study to a large extent further builds our African collaborators' capacity in a big team science replication. African collaborators participated in instrument cultural adaptation and translation, ethics application, local site coordination, study nomination and assessment, and data collection. All these contributions provided on-the-job training to our African and even to some of our non-African collaborators. As a result, 57 African researchers are now more familiar and have built capacity in some crucial tasks required for contributions to big team science.

Difficulties. We faced a number of challenges when conducting the ManyLabs Africa study that bear mention to assist others wishing to conduct similar projects.

1. **Poor infrastructure.** This challenge was peculiar to our local African sites and we couldn't move forward with some. Infrastructure and resources are scarce for African labs and collaborators which hampered their speed and size of sample collected.
2. **Sites/Collaborators non responsive and withdrawal.** Although we had 264 sites that showed interest in collaborating on this study, only 43 were able to contribute. A vast number of them were not responsive while some withdrew their interest due to their heavy workload and impractical schedule. While a couple of labs withdrew due to ethical rejections. Perhaps most European and North American labs/collaborators withdrew or were disinterested in the replication because the studies selected were not

relevant and of priority to their research institutions or them. However, most of our European and North Americans who collaborated considered the African-origin effects applicable and important, fifteen of which applied and obtained ethics approval from their local reviews, and were committed and contributed extensively to the study.

3. **Study search.** Searching African literature that meets our preset criteria was difficult. African collaborators nomination was helpful but most studies we found used qualitative (interview and focus group) designs, Western-developed instruments, and to an extent written in Arabic and French, languages the lead author has not mastered. The studies we could choose from were limited and it took about 6 months to complete the entire study selection. The full report of our study search and selection is available at <https://osf.io/v6j28>.

We elaborate further on the capacity building aspect and our difficulties/realities of conducting big team science in Africa in detail in Chapter 6.

Our replication compares to others. The ManyLabs Africa study approach is different from other big team science including our CREP Africa study in a number of ways. Unlike other big team science studies that mostly investigate Western-origin effects, ManyLabs Africa is the first study that prioritized African claims and generalized these claims from Africa to other African, European, and North American populations. The idea was to foreground African thoughts and expand African theories to these populations. The generalization of extramarital sexual relationship claims across our African, European, and North American participants made a point of how African claims can expand what we know of other cultures. We believe this replication could further stimulate discourse on why/why not African and non-African scholars should generalize from Africa to global populations as well as provide an example on replicating African claims on a big team science scale.

Despite the infrastructure and resource challenges encountered by our African sites such as lack of computer, intermittent and expensive internet and power, no participant compensation, that limit the number of African sites we worked with, this study was able to increase participation of African participants (1136/6009, 18.9%) and collaborators (36/115, 31.3%) compared to, for example 12.5% African participants and 2.7% African researchers involved in the Psychological Science Accelerator COVID-19 studies (Buchanan et al, 2023).

Unlike other big team science studies that relied on convenient claims that were hand pick or proposed or nominated from expert or peers for replication, the ManyLabs Africa adopted a 3-step bottom-up study selection approach that engaged African collaborators to nominate actual paper and study focus and also surveyed all local collaborators from the target study populations to evaluate and assess shortlisted studies. This workflow ensured that our replications were of interest, relevance, ethically acceptable, feasible, adaptable, applicable, and priority first to Africans and to all target study populations.

10.2 Limitations

We do not intend to generalize the current study findings to the entire continents because Africa and the other continents sampled are extremely heterogeneous and our samples missed much of this heterogeneity. For instance, our sample includes only 5 of 54 African countries, 11 out of 44 European countries, and 2 out of 23 North American countries. Also, our participants are mostly urban dwellers and internet users. Moreso, internet penetration in 2 out of the 5 African countries (Tanzania, 31.9%; and Burkina Faso, 19.9%) where our African sites are located is low and below the African average (43%; Galal, 2024), and where available mostly affordable by the privileged. Our African labs also lacked the funds to take our experiment to rural areas, resulting in the urban skew of our sample. Lastly, our samples were relatively young ($M_{age} = 26.41$). This particularly was noticeable when we applied the inclusion

criteria of parents aged 35 and above for the teenage pregnancy replication focal subgroup. Only 695 out of the full sample of 5806 met this inclusion criteria. While we were still able to include Europe (n = 558) and African (n = 113), the resulting sample of over-35 parents in North America was so small that it had to be excluded entirely (n=24).

10.3 Recommendations

The Manylabs Africa study has enriched the sexual behavior literature, big team science, especially regarding replication from an African perspective and helped build our African collaborators' capacity. Yet the low replication and generalizability of these African claims, and the realities of conducting big team science in Africa underscores the need for more global investigations for generalizability and applicability. With increased focus on cross-cultural multinational studies (Jarke et al., 2022), it is increasingly important that African researchers (and of course their colleagues across the globe) adopt research best practices such as study preregistration and sharing accessible and usable research materials (Azouaghe et al. 2022) to produce research that is reproducible and credible. Researchers interested in generalizing our findings beyond these replication samples may adopt our materials and workflows. Perhaps worthy of note is our study selection workflow that allows project contributors to provide vital inputs at the conceptualization stages and to identify study populations' priorities. The realities of conducting research in low and middle income countries require that researchers and funders carefully consider available resources and facilities (Adetula et al., 2021; Emery et al., 2023; Onie 2020).

10.4 Conclusion

Investigating claims generalizability is crucial to understanding humans particularly to the relatively unknown and underrepresented Africa. Our sample African and North American

samples lacked the diversity that exists on these continents, most participants from these continents were non-parents younger than 35 years. Overall, this study showed that African claims on attitudes toward teenage pregnancy, condom use, and extramarital affairs to a small extent can generalize across and beyond Africa. and contributes evidence for cultural homogeneity within African, European, and North American populations for these behaviors, at least in the urban, educated, internet-literate subsamples that we could reach in our target countries. This study provides additional context to the credibility and applicability of psychology claims discourse. It is important to continue to investigate psychology claims across Global North and Global South. There is more to do to increase African thoughts and participation globally. In the next chapter, we reflect more generally on the connections between the credibility revolution and global development.

Chapter 5: Synergy Between the Credibility Revolution and Human Development in Africa

Psychological science is at a crossroads. Practices once deemed rigorous now appear to lead to non-replicable findings. The ensuing “credibility crisis” (Vazire, 2018) has led to a period of intense scrutiny, mostly among North American and European scientists, of problems in science and how to fix those problems. These initiatives aim to increase the accessibility of science, provide free digital and intellectual infrastructure to support rigorous science, develop practices that lead to credible evidence, and make science more collaborative. This current thesis used the tools, workflows, and scholarship from these initiatives, both to conduct the CREP Africa and ManyLabs Africa replications and to build capacity among our African collaborators to do open science and big team science.

However, this dissertation is one of a small number of initiatives that is leveraging the learnings of the credibility crisis to do better science in lower resource settings. Nor have activists in the credibility crisis grappled substantially with problems that affect the credibility of psychological research, but not via poor replicability, such as the overreliance on convenience samples (Arnett, 2009; Henrich et al., 2010; Tiokhin et al., 2019) and the limited applicability of psychology to applied problems (Premachandra & Lewis, 2022). Even, our failure to replicate and generalize 14 out of the 18 African claims to Europeans and North Americans further raises questions about method, settings, and cultural variations that may exist in Africa or whether African psychology also suffers from poor research quality and questionable research practices that prompted the credibility revolution. In any case, addressing this problem requires building capacity and good working conditions which most African countries lack, for a more robust and inclusive psychology.

African science is at a crossroads of its own. In the wake of the devastation wrought by colonialism, African nations have struggled to build the infrastructure and institutions necessary to support a happier, healthier, and more educated population. International organizations such as the UN also increasingly recognize African development as a broadly important goal; with 17% of the world's population that should increase to 42% by century's end (Roser et al., 2013), the world cannot realize important trans-national goals, such as the UN's Sustainable Development Goals (<https://sdgs.un.org/goals>), without engaging Africa. Research and post-secondary education can assist with this development through conducting robust research to help with evidence-based policy, building human capital, and assisting with technological catch-up (Atickem et al., 2019; Bloom et al., 2014).

Yet acquiring higher education and conducting research in Africa have lagged due to problems that include low funding, inadequate infrastructure, and political instability (Kumwenda et al., 2017; Nabyonga-Orem et al., 2020; Nsamenang, 2007; Ojiji, 2015). These education and research challenges are reflected in the slow development and poor impact of psychology, a research field that Cameroonian psychologist Bame Nsamenang (2007) described as “nascent” in most parts of sub-Saharan Africa (see Ibrahim, 2011, for similar arguments about the Arab world). While psychology is not strong in Africa, neighboring disciplines such as public health, development economics are gaining traction, especially applied behavioral science (Hallsworth, 2023). For instance, WHO organized a regional workshop in Eastern Africa adopting a Training of Trainers model to build participants' capacity in theories, methods, approaches and evidence in behavioral insights (World Health Organization, 2022).

Development communities, scientific research institutions, and research psychologists globally, especially those in Africa, have the huge challenge of educating and conducting research to inform policies and provide more psychological services to the African people. We

believe that there is synergy between the credibility revolution in psychology and this goal. This synergy occurs because of the credibility revolution's emphasis on building free, accessible, and credible processes and tools to improve all parts of the research process (see Korbmacher et al., 2023 for further discussion). These processes and tools have the potential to accelerate the development of the human capital and infrastructure necessary to build stronger, more vibrant research communities throughout the continent (Mwangi et al., 2021; Onie, 2020; UNESCO, 2021). If the credibility revolution's newly-developed processes and credibility-enhancing tools are applied to concrete problems that accelerate human development, psychology could even reach a goal it has long aspired to, namely using its research to directly assist in solving big social problems such as corruption and fraud, social and economic inequalities, and ethnic conflicts. Psychology also stands to benefit from stronger psychology research communities in Africa; not only because these stronger communities could lead the research initiatives that help make psychological findings more generalizable to Africans, but also because such communities could push psychology to study issues currently outside its mainstream.

We attempt to build bridges between the credibility revolution and the movement to increase human development in Africa. This piece describes the current state of African psychology research and how decision-makers in these world regions can facilitate the spread of open science practices in Africa. It provides an overview of the credibility revolution and how the new open science innovations that developed in response to the credibility crisis can support credible and accessible research.

1. The research landscape in Africa

Africa is an enormously diverse continent, home to 1.3 billion people from speaking over 1000 languages across 54 countries. The amount of ethnic diversity in Africa is also high:

in sub-Saharan Africa (the region that excludes Africa's North coast), there are about 35% more ethnic groups per country than any other world region (Fearon, 2003, Table 1). Finally, African nations are each subject to their own unique political circumstances, which deeply impact government support for research and higher education. These factors complicate any attempt to clearly summarize Africa's research landscape.

What does unite the African continent is a shared history of subjugation to colonizing powers. Several decades of colonial regimes and economic resource exploitation by colonizing powers (Michalopoulos & Papaioannou, 2020) left a legacy of political and economic instability that has obstructed the development of strong academic research institutions (Nsamenang, 2007). In the post-colonial period of the 1960s, the United Nations made a concerted push for the newly-independent African nations to establish systems of secondary and tertiary education. In the 1970s, these priorities shifted to universal primary education, with tertiary education being largely neglected (Bloom et al., 2014). University-housed research communities in most regions of Africa suffered (Bloom et al., 2014), stymied by high workloads, poor IT infrastructure and an unstable power supply, burdensome administrative requirements, inadequate funding, and corruption (Atickem et al., 2019; Nsamenang, 2007; Ojiji, 2015). Particularly telling are those of psychology and behavioral science, which are relatively young, poorly funded, and emphasize more on directly contributing to economic development rather than research scholarship (see Mughogho et al., 2023; Kenmoe, 2022 for context).

Today, the political situations in most African countries have stabilized and most economies are growing, sometimes quite rapidly (Young, 2012). Higher education is following suit: enrolment in tertiary education in Africa grew by 170% between 1999 and 2002 (ICEF Monitor, 2015). In Malawi for example, there are efforts to improve higher education enrolment (Sharra, 2023). Some international efforts are also attempting to fund and lead

policy-informed multidisciplinary research and capacity-building programs to improve the scholarly ecosystem across Africa to meet human needs. To mention a few, the World Bank's Zanzibar Improving Quality of Basic Education Project (<https://projects.worldbank.org/en/projects-operations/project-detail/P178157>) aims "to improve teaching competencies and learning outcomes and reduce the gender gap in transition rates within basic education in Tanzania." In another example, the Bill and Melinda Gates Foundation (<https://www.gatesfoundation.org/our-work/places/africa>) has a global education program, in which they engage with a variety of partners, including local and sub-national governments, to build capacity in the education sector. In addition to these large, well-funded organizations, smaller organizations are also attempting to build capacity in Africa's scholarly ecosystem. For example, the Center for Effective Global Action (CEGA; <https://cega.berkeley.edu/>), a global network of researchers, has a "Global Networks" program aimed at supporting the scholarly ecosystem in low- and middle-income countries, especially those in Africa.

Despite this growth and efforts, an enormous gap in research capacity persists (Atickem et al., 2019; Musau, 2017). To understand the scale of this gap, consider that higher economic output reflects, in part, an expansion in the available goods and services available to the average person. The average GDP per capita in the United States was \$54,225, about 15 times the value of \$3,467 for sub-Saharan Africa (Roser, 2013). To make this difference concrete, our US-based readers might imagine that you have a pool of \$2,000 for a particular research project but everything about that research – from participant recruitment costs, to data analysis software, to publishing charges, to fees for review by Institutional Review Boards – costs 15 times what you currently pay for it. Moreover, due to poor institutional resources, the \$2,000 might need to come out of your own personal funds, where it must compete with the money needed to cover your basic living expenses (Atickem et al., 2019).

In this context, it is no wonder that discussions of “building research capacity” in Africa often prioritize research areas such as psychology and behavioral science that can directly resolve concrete problems for Africans, such as maternal health, poverty, and infectious diseases (Nordling, 2018; The World Bank Group & Elsevier, 2014). Small wonder, as well, that African researchers tend to prioritize and study applied rather than basic research (Hountondji, 1990). Psychology may need to demonstrate its direct applicability to African human development problems and to maximize the benefits of the scholarly ecosystem on the continent if it wishes to attract the attention of talented prospective African researchers and funders respectively.

2. The credibility revolution in North America and Europe

In 2011, a well-regarded social psychologist, Daryl Bem, published an unusual article in the world’s top social psychology journal, the *Journal of Personality and Social Psychology*. This article was unusual because of its central claim: that undergraduate students who he had recruited from Cornell University in Ithaca, New York, USA could be affected by events that hadn’t yet occurred (Bem, 2011). In other words, these undergraduates could sense future events. What’s more, Bem applied experimental design and statistical methods that met or exceeded social psychology’s standards of rigor at that time. This left one of two possible explanations: either everything we knew about physics was wrong or those methods were not as rigorous as they seemed (see Francis, 2012; Schimmack, 2012; Shrout & Rodgers, 2018).

This realization spurred a trend of large replication studies designed to examine the credibility of past findings. These replication studies took a protocol that had been previously used to claim a particular discovery, then recruited new participants with that protocol – often a much greater number than the original paper achieved. The replicators were able to achieve

these large sample sizes by forming very large collaborations across many labs, thereby allowing them to investigate a past finding with much greater precision than was feasible for any lab on its own (Forscher et al., 2022). The findings of these collaborative replication studies were worrying; across 77 large replication studies, 43 reported statistically significant results in the same direction as the original, with effect sizes that were only 53% as large as the original (Nosek et al., 2021). Although there is no consensus about what the replication rate of psychological research ought to be, it is at least clear that current methods and practices led to a literature that substantially overstated the evidence behind many findings (Spellman, 2015; Vazire, 2017). The task, then, was to identify problems that undermine research credibility and identify practices to fix those problems.

The result was an explosion of research on practices to improve the credibility of psychology research. Most of these practices focus on bringing something into the open that was previously hard to access. For this reason, we call them “open science practices”, though this broad label obscures major differences in goals and motivations (Fecher & Friesike, 2014). We believe these practices have the potential to improve many aspects of psychological research. However, much of the focus of the credibility revolution has been on addressing the problem that kicked off the revolution, namely poor replicability of findings. For this reason, the innovations of the credibility revolution have yet to be fully applied to problems outside replicability. Most replication studies, for example, focus on the replication of low-cost studies that can be completed online with convenience samples drawn from college classrooms or platforms like mTurk (Gervais, 2021). This feature means that, even if the credibility of the evidence improves as a result of the new replication studies, the findings still have limited applicability to large swathes of humanity and their problems.

3. How open science practices can contribute to human development

Although the open science practices used in psychology’s credibility revolution have largely been applied to improve research credibility via better replicability, they can in principle be applied to address a broad range of problems in a broad range of topics.

Herein lies the potential for cross-fertilization between the US and Europe’s “credibility revolution” and efforts to expand human development in Africa. We discuss some of the relevant tools that can help improve African research accessibility, evidence quality, and infrastructures. By no means are the tools mentioned exhaustive as credibility revolutionists continue to advance these tools.

Table 18 Five categories of open science practice and initiatives that fit each category.

Category	Name	Link	Purpose
Accessibility	PsyArXiv	https://psyarxiv.com/	Freely share and access preprints related to psychology
	AfricArXiv	https://info.africarxiv.org/	Freely share and access preprints related to Africa
	MERLOT	https://info.merlot.org/	Freely share and access teaching materials
	Sci-Hub	https://en.wikipedia.org/wiki/Sci-Hub	Freely and illegally access journal research articles
Infrastructure	Open Science Framework	https://osf.io/	Freely share and access research products, including study materials, research code, datasets, and protocols
	Open Researcher and Contributor ID	https://orcid.org/	Freely create a profile and unique open digital identifier for a researcher and share publications.
	Github	https://github.com/	Freely share and access research code
	Dataverse	https://dataverse.org/	Freely share and access research data
	FigShare	https://figshare.com/	Freely share and access up to 5GB of space
	PsychoPy	https://www.psychopy.org/	Freely program research experiments
	R	https://www.r-project.org/	Freely program analysis code
Credibility	RStudio	https://www.rstudio.com/	Free, user-friendly interface for R
	Registered Reports	https://osf.io/rr/	Get journal peer review and approval before you collect research data
	Peer Community In Registered Reports	https://peercommunityin.org/	Get (scheduled) peer review and approval before you collect research data, and recommendations to over 30

			PCI-friendly journals and other journals for publication
	OSF Preregistrations	https://osf.io/prereg/	Store a detailed study protocol before you collect research data
	AsPredicted	https://aspredicted.org/	Store a detailed study protocol before you collect research data
Community	The Psychological Science Accelerator	http://psysciacc.org	A large, standing network that conducts big team science psychology projects on all five populated continents
	Collaborative Replications and Education Project	https://osf.io/wfc6u/	A large, standing network that uses big team science to train bachelor- and master-level students
	ManyBabies	https://manybabies.github.io/	A large, standing network that conducts big team science projects focused on infancy research
	StudySwap	https://osf.io/meetings/studyswap/	A platform to develop collaborations and exchange resources across labs
	African Reproducibility Network	https://africanrn.org/	A continent-wide reproducibility network to improve and advocate for open science practices and research reproducibility across Africa
	Framework for Open and Reproducible Research Training	https://forrt.org/	A pedagogical community to advance research transparency, reproducibility, rigor, and ethics via pedagogical reform and meta-scientific research.
	ReproducibiliTea	https://reproducibilitea.org/	A platform to set up and conduct journal clubs at local institutions

This potential exists because, although the open science practices are both domain-general and varied, they all aim to open many different aspects of the research process. For the purposes of this article, we separate the practices into four groups according to the goal they aim to achieve. Initiatives to increase *research accessibility* allow researchers to freely share and access written research articles. Initiatives to provide *infrastructure* aim to provide free resources to make the research process easier. Initiatives to improve *credibility* aim to improve the quality of research evidence. Finally, *community* initiatives aim to build communities to accomplish specific, often open-science-related goals.

The *accessibility* initiatives will perhaps be most familiar to an African audience. These initiatives note that, after research is published as an article in a scientific journal, the publisher takes the article copyright and locks the article behind a paywall. From the perspective of

promoting research credibility, these paywalls are a problem because they restrict access to information that is needed to inspect evidence credibility. However, paywalls are also a problem from an equity perspective because they restrict the access of the very people who might most benefit from scientific knowledge and innovation.

Most Africans will be familiar with one solution to this problem (Huang et al., 2020): a journal business model that allows the article to be accessed for free, as long as the submitting author pays an “article processing fee” to the journal. Under these *open access business models*, the journal keeps the article copyright but the journal imposes an additional fee on article submitters. Unfortunately, open access business models are not ideal for article authors at low-resource institutions. This is because the business model shifts the costs of maintaining digital versions of an article onto authors, who may not be able to pay this fee if they are at an institution with low financial resources. Some open access journals can even be “predatory” in the sense that they provide publications to researchers in exchange for a fee while doing little to ensure that their published papers have scientific merit (Mouton & Valentine, 2017; Xia et al., 2015). Not all open access journals are predatory in this sense; journals are “predatory” only if they are strictly money-making enterprises that provide little to no quality control, with the result that they allow the publication of pseudo-scientific or nonsense articles because the authors of these articles nevertheless paid a publication fee.

Beyond open access journals, the credibility revolution has popularized some additional initiatives that are not journal-focused. Many of these initiatives focus on increasing the accessibility of a version of the article where authors keep the copyright: the version just after article acceptance, but before the journal does any formatting or copy-editing of the accepted article, often called the *preprint* (for a comprehensive guide, see Moshontz et al., 2021). Authors have the right to share this version of the article. Thus, a variety of websites have sprung up to host this version of the article, such as PsyArXiv and AfricArXiv. These websites

also assign Digital Object Identifiers (DOIs) to the articles they host, which allow the preprints to be indexed by databases, such as the database maintained by Google Scholar. Journals often have special policies that govern how they handle preprints; authors should consult those policies before they post preprints to these servers. The Sherpa database provides a searchable database of those policies (<https://v2.sherpa.ac.uk/romeo/>).

One final accessibility initiative deserves mention. The article database Sci-Hub provides access to the published versions of articles (post copy-editing and journal formatting) entirely for free. However, the use of Sci-Hub is illegal in most countries. Thus, we cannot recommend the use of this service – particularly for African researchers, whose libraries may not be able to afford the expensive fees that most journals charge.

The *infrastructure* initiatives aim to make a certain aspect of the research process easier by providing free infrastructure to enhance the research process. For example, the Open Science Framework (<https://osf.io>) is a general-purpose website created and maintained by the Center for Open Science that allows researchers to archive and post their research materials, data, analysis scripts, and other research artefacts. At any time, a researcher may choose to make these materials public, in which case they may be accessed by others for free. This can benefit both researchers who access the open materials, because they can re-use the materials and can help improve processes and workflow for their own research, and the researcher who shares their materials, because their reputation and credibility are enhanced (Ebersole et al., 2016). In addition to the general-purpose Open Science Framework, there are many specialty initiatives that focus on allowing the free sharing of one type of research product, such as research code (Github; <https://github.com/>) or research data (Dataverse; <https://dataverse.org/>).

This category also includes many free and open source computing languages designed to facilitate research computing. This includes the PsychoPy language, a language to facilitate the programming of psychology experiments (e.g., reaction to expressions, reaction time), and

R, a language to facilitate the analysis of research data. These computing languages often take time to learn, but there is no financial cost to using them. In addition, a great deal of tutorial materials exist for learning these languages. For a tutorial for PsychoPy, see <https://www.psychopy.org/coder/tutorial1.html>; for a tutorial for R, download the free R-based programming environment, RStudio, and see <https://r4ds.had.co.nz/>.

The *credibility* initiatives consist of workflows and policy frameworks that help ensure research findings are transparent and credible. The most prominent of these are Registered Reports, a framework for a research article format in which authors submit research protocols that are peer reviewed before any data collection occurs. If the protocol is deemed sufficiently rigorous during the peer review process, it gains “in principle acceptance”: as long as the researchers follow the protocol during the collection and analysis of data, the paper is accepted, regardless of how the data turn out. This policy framework helps protect against the selective publication of positive results, a practice that is prevalent in the social sciences and that undermines the credibility of entire research fields (Scheel et al., 2020). Many journals already support the Registered Report format, and their number continues to grow; for a list, see <https://cos.io/rr>.

A second type of credibility initiative is websites for *pre-registering* a study protocol. Essentially, a researcher prepares the first part of a Registered Report and archives the protocol on one of these websites prior to running the study. This pre-written protocol allows outside researchers to evaluate the match between the protocol and the final article and check whether deviations from protocol are justified. Under a mode of science that prioritizes falsifying pre-specified predictions, preregistration increases the credibility of the final evidence (Kaplan & Irvin, 2015; Lakens, 2019). Preregistration does not necessarily involve any formal peer review of the written protocols and is not tied to any specific journal. Thus, unlike a Registered Report, a researcher who preregisters one of their studies is not limited in their choice of journals.

Finally, the *community* initiatives aim to build communities to accomplish specific goals. A few of these initiatives are aimed at increasing social connections between members. One of these is ReproducibiliTea, a platform for setting up journal clubs at your own institution. However, some of these networks are more focused on completing very large, often international, collaborative projects. These “big team science” initiatives (Forscher et al., 2022) play a central role in “scaling up” the amount of resources that are invested into replication studies so that claims can be carefully examined. However, big team science can also deploy large numbers of people on a new study of central importance. For example, one prominent big team science organization, the Psychological Science Accelerator, initiated a project to investigate the psychological aspects of COVID-19 (Buchanan et al., 2023). By pooling resources across the large collective, this project was able to recruit over 47,000 participants from 110 countries, using materials in 44 languages for three primary research studies. Big team science organizations also exist for infancy research (the *ManyBabies* initiative; Byers-Heinlein et al., 2020) and for teaching undergraduates the research process via replication studies (the *Collaborative Replication and Education Project* or *CREP* (Wagge, Baciú, et al., 2019; Wagge, Brandt, et al., 2019). Readers will find a variety of materials useful for integrating the CREP model of replication research into undergraduate classes, including links to tutorial videos subtitled in a variety of African languages, at <https://osf.io/4v5yg/>

As of yet, big team science organizations do not have a strong presence of African researchers. For example, as of 2022, only about 2.07% of the membership of the Psychological Science Accelerator was from Africa (Kijilian et al., 2022). We believe this lack of African members represents both a gap and an opportunity: a gap because these organizations need a stronger African presence if they are to be truly international, and an opportunity because African researchers who join these organizations can become integrated into broad networks dedicated to impactful research.

4. Broadening the open science tent in Africa

Even though we believe open science practices can facilitate the development of the capacity to conduct research in Africa, they are not themselves sufficient due to lingering barriers to African research capacity. Fully addressing these barriers will require concerted effort on multiple fronts from the many stakeholders, African and non-African, that comprise the scientific ecosystem. Here we describe some of these steps these stakeholders can take to facilitate and enhance the impact of the open science practices from the previous section. Our suggested steps fall into four categories: *resources*, *training*, *research tools*, and *ethics* (see Table 19).

Table 19 Barriers to expanded research capacity and interventions to address those barriers

Category	Barrier	Interventions	Stakeholders				
			U	S	F	J	R
Resources	Financial resources in Africa are limited	Provide small, unrestricted grants with few accounting requirements to African researchers	X	X	X		
		Provide automatic waivers for Article Processing Charges that overstretch limited financial resources for African researchers				X	
	Infrastructure and facilities are limited and unreliable	Provide internet grants to allow African researchers to purchase better internet-based resources	X	X	X		
Training	Training is needed to use open science practices	Create a train-the-trainer program for African researchers, which can focus on encouraging interactive online tutorials, lectures, workshops and provide certificates of participation	X	X			X
		Organize African research-focused training events such as workshops, seminars, and talks on how to access and adopt cutting-edge research resources and practices	X	X	X		X
		Establish, support, and fund open science labs in Africa, which can serve as designated research centers to advance open science practices	X	X	X		X
		Participate, collaborate, support and fund African-based open science capacity building initiatives such as the Training Centre in Communication Africa, Eider Africa	X	X	X	X	X
Research tools	Measures of psychological processes developed in the US and EU may not generalize to Africa	Conduct exploratory research to adapt existing measures and develop new ones			X		X
		Create easily-accessible databases of measures and their translations or adaptations, along with evidence about their validity in different populations			X		X

	Many existing tools are exclusively in English	Create and fund translations of existing tools into different languages, including the indigenous languages of Africa				X		X
	Mainstream research questions may not be relevant to Africans	Conduct research on practical problems in Africa						X
		Actively engage African researchers in the early stages of the research process						X
Ethics	Predatory journals let African researchers engage in a pay-to-play publishing schemes	Create African open access overlay journals on preprint servers that will lower the incentive to publish in predatory journals	X	X	X	X		
	Institutionalized exploitative practices rob African researchers of funds and exacerbate working conditions	Establish research integrity offices led by African researchers and morally and financially supported by North American and European researchers	X					X
	Non-African researchers practice research colonization by using African researchers as data collectors without giving adequate credit or compensation	Adopt contributorship systems to document fine-grained contributions to research					X	X
		Use collaboration agreements at the start of projects to set explicit expectations about expected contributions and systems of credit and compensation			X			X

Note. U = Universities; S = Scientific Societies; F = Funders; J = Journals; R = Researchers.

The first category of barriers, a lack of financial and non-financial *resources*, is the primary barrier to additional research infrastructure that is mitigated by the use of open science practices. This is because open science practices tend to emphasize the use of free or open-source tools to make up for the lack of pre-existing infrastructure. Yet these open science practices cannot by themselves compensate for the resource restrictions that many African scientists face. For example, many African scientists lack reliable internet, which is a prerequisite for using most of the open science resources we previously described; thus, internet grants can make the use of free open science tools more feasible. In the era of online conferences, these grants can also greatly facilitate the participation of African scientists in the global academic commons. Restricted research funding means that, for many African scientists, even small grants can be instrumental for completing research projects – as long as these grants are not saddled with burdensome accounting restrictions that make the money too much of a hassle to be worth using. Although we have emphasized the use of preprint servers

rather than open access journals, those journals that do use an open access model can address resource-related barriers by providing automatic waivers (with few bureaucratic barriers) of article processing charges to researchers from low-income countries (Nabyonga-Orem et al., 2020). Finally, resources should go to existing initiatives that are on the continent, which are numerous but often lack resources. These initiatives include among others, the Training Centre in Communication (TCC Africa; <https://www.tcc-africa.org/>) that advocates for Global North-South higher level engagement on open science policy, builds open infrastructure for data access and sharing, and provides effective communication skills to improve scholarly output and visibility. LIBSENSE (<https://libsense.ren.africa/en/>) is a collaborative initiative of African research and education networks led by the West and Central African Research and Education Network (WACREN, <https://wacren.net/en/>), and regional African research and education networks (ASREN, <https://asrenorg.net/> ; UbuntuNet Alliance, <https://ubuntunet.net/>) to advance open science and open access practices in Africa. African Reproducibility Network (The AREN team, 2023; <https://africanrn.org/>) is a “continent-wide reproducibility network to improve and advocate for open science practices and research reproducibility across Africa.”

Barriers related to *training* refer to the fact that, even though open science practices often rely on freely-accessible infrastructure, using this infrastructure requires specialist training. At first, some of this training may need to come from North American and European sources due to the better-developed research cultures and infrastructures in these world regions. However, we believe training activities should emphasize establishing grassroots reservoirs of knowledge and skills in Africa so that, eventually, most training comes from local rather than foreign sources. Thus, in addition to activities such as African-focused workshops, training activities should emphasize train-the-trainer programs and Africa-based “open science labs” that serve as training centers for open science practices. Local African educational institutions should incorporate open science content in research methods courses for more enduring

benefits and improvements of these practices. We should also encourage and learn from existing initiatives to advance capacity in open science practices and tools in Africa such as Eider Africa (<https://eiderafricaltd.org/>), an organization that trains African researchers via mentorship initiatives.

A third category of barriers are those related to *research tools* that may not be adapted for African realities. For example, most research tools are designed and implemented in English. Although many people in African countries speak English, Africa is highly multi-lingual. Translating these research tools to other languages using the services of professional translators or at a minimum (because translation software could be language-limited or inaccurate), automatic translation software such as that provided by DeepL or Google Translate, is therefore one way to ensure that more Africans are able to use these research tools. A more sustainable approach to producing African-friendly research tools is to facilitate cooperation between the developers and local users to factor in African dynamics when creating these tools.

However, the research tool barriers are not merely linguistic. Existing measures of psychological processes may not even be applicable in a specific African culture. Such measures may need extensive cultural adaptation, and the target psychological process may not even exist at all. Facilitating such time-consuming measurement work will also require a comprehensive database of psychological measures, their adaptations, and evidence about their validity in different cultures. In addition, some psychological processes may exist in Africa that have no clear equivalents in North America or Europe. Hence co-designing measures rather than validating measures designed in one population for another could produce more generalizable measures. Descriptive, qualitative, and exploratory approaches that actively involve African researchers at all stages are necessary to identify and understand these

processes in their original context. This descriptive work contrasts with approaches that take North American and European as a baseline against which African psychology is compared.

Even more radically, psychology as it is currently constituted may simply study topics that are uninteresting or irrelevant to stakeholders in Africa. To make the case that psychology research is worth the investment of Africa's limited education funds, psychology may need to put a more central focus on studying problems directly relevant to Africa's ongoing concerns (Lakati & Masibo, 2023). This may mean focusing on solving practical problems, such as those related to maternal and mental health issues, road accidents, and the transmission of infectious disease.

Trans-national benchmarks such as the UNESCO's Operational Strategy for Priority Africa 2022-2029 (UNESCO, 2023) which aligns with our advocacy for enhancing open science practices and credibility enhancing-tools for research and development in Africa may provide a useful guide for the sorts of topics that can directly assist with African development. This UNESCO strategy focuses on African-driven strategy intervention to strengthen higher education, improve digital skills and competencies and build capacity to increase research, innovation and technology production via effective networking, open science and best research practices. There are country-level African priorities, such as Kenya's reproductive health research priority 2022-2027 (Hearty, 2022), the South African health department's National Health Research Strategy: Research Priorities for South Africa 2021-2024 (<https://www.health.gov.za/wp-content/uploads/2021/04/NATIONAL-HEALTH-RESEARCH-STRATEGY-2021-2024.pdf>), and the national research priorities from the Tanzania Commission for Science and Technology (<https://www.costech.or.tz/ResearchPriorities>), that could be pivotal for attaining research credibility and human development.

The final category of barrier, *ethics*, refers to ethical pitfalls that can short-circuit the development of robust psychology research communities in Africa. The dominance of North America research cultures and theories of global psychology is well-documented (IJzerman et al., 2021; Thalmayer et al., 2021). The overwhelming influence of Western-developed perspective and approach to explaining an African phenomenon does not only raise questionable applicability to Africans, but also ethical concerns. Some scholars argue that this overshadowing North American influence has hampered the development of indigenous African psychology, and hence there is a need for decolonization of African psychology (Oppong, 2022). Even North-South research partnerships create yet another problem, North-South power imbalances (Dannecker, 2022; Green, 2021; Matenga et al., 2019). These power imbalances create the risk research neo-colonization (Boshoff, 2009), in which African research partners are treated as “data collectors” for a foreign collaborator benefit and are given insufficient credit and funds for their contributions (Van der Veken et al., 2017). In neo-colonial arrangements, the party with the money dominates the research questions or scope of the research with little consideration for African interests and priorities that are needed for local impacts (Dannecker, 2022; Kok et al., 2017) .

Part of what complicates this issue is that, due in part to the highly varying levels of development across countries, different researchers may enter collaborations with wildly different expectations about what constitutes “sufficient” compensation. We believe the adoption of *contributorship systems* (Holcombe, 2019), in which authors explicitly document fine-grained contributions to research papers in, for example, the article author notes, can at least make visible when research colonization is occurring. As a side benefit, more robust systems of contributorship could also illuminate the problem of guest authorship by powerful researchers. Explicit *collaboration agreements* that lay out expected project contributions and systems of compensation ahead of time can also help align the expectations of all parties in a

collaboration, thereby mitigating the risk that one party feels exploited (see Forscher et al., 2023).

A second ethical pitfall relates to an issue we mentioned previously, the existence of predatory journals that take advantage of pressures to hit publication metrics by offering publications in exchange for a fee, with little regard for the quality of the finished research. We believe these pressures to publish can be redirected into more productive avenues by leveraging the preprint servers that have developed as part of the credibility revolution's accessibility initiatives. The preprint servers provide a place to deposit papers; one can turn these preprints into publications by developing so-called "overlay journals" that peer review and curate, at no charge, high-quality papers that fit a particular theme on a preprint server (Brown, 2010; Smith, 2000). Because overlay journals rely primarily on the IT infrastructure used to create a preprint server, their costs tend to be lower than the costs of a traditional journal. The costs to support these overlay journals should therefore be within reach of existing scientific societies. The existence of these journals should provide competitive alternatives to African authors wishing to publish their research.

A third pitfall relates to the existence of institutionalized practices that exploit African researchers, such as the misappropriation of research funds (Ojiji, 2015) or guest authorship by researchers in positions of power (Rohwer et al., 2017). These practices are certainly not restricted to Africa, but the political instability in the wake of de-colonization has entrenched such practices in some African countries (Kirya, 2019; Lawal, 2007). Such entrenched practices may require the establishment of institutional research integrity offices such as African Research Integrity Network (<https://africarinetwork.wixsite.com/website>) that are led by Africans but should be supported, morally and financially, by North American and European researchers. These African research integrity offices should be non-profit, independent, and

internationally affiliated organizations with a mission to improve research credibility, and research fund management and performance.

5. Conclusion

North American and European psychology and African science are both at pivotal moments where they have the potential to make lasting improvements. This potential exists because of the credibility revolution on the one hand and the push for human development on the other. We believe these two movements can benefit each other. The credibility revolution can benefit through engaging researchers of different perspectives to solve problems that concern a greater slice of humanity. African science can benefit from open science practices that use open infrastructure and enhance research credibility. We hope both these movements take advantage of this crossroads to change themselves for the better.

Chapter 6: General Discussion, Conclusion and Recommendations

The CREP Africa and Manylabs Africa studies are innovative. For example, this is the first rollout of a CREP study in Africa thereby providing our African collaborators the opportunity to improve their research capacity in open science practices, and also to investigate a moral judgment theory among some Africans. The ManyLabs Africa is one the few studies that replicated African claims across Africa, Europe, and North America on a big-team science scale. These studies introduced some research workflows such as: the bottom-up study selection for a replication study, and the double-forward translation procedure (Adetula et al. 2021) that could help improve research workflows. Owing to lessons learnt from these studies, we summarize below our takeaways and recommend practical solutions to researchers interested in conducting big team science investigations in Africa and broadly to research stakeholders on the continent and beyond.

1. Need for more scientists and capacity building

While we don't claim there are no good and seasoned researchers in Africa, there is a need for more African scientists and for capacity building (Bryant, 2014; Kumwenda et al., 2017) especially with the on-going credibility revolution that has necessitated the adoption of cutting edge open science practices and methods (Azouaghe et al. 2021; Adetula et al., 2022). Across these two studies, 23 African collaborators have attended open science workshops, nominated studies/articles and assessed the shortlisted studies for the final selection, culturally adapted and translated the measure from English into five African languages (Arabic, Chichewa, French, Swahili, and Portuguese), obtained local ethics clearances, coordinated their local labs to collect data, read and reviewed the studies' manuscripts, and managed a page on

the Open Science Framework website that curates their site-specific materials. We believe these activities have built capacity among our collaborators and equipped them with appropriate expertise and networks for further contributions. Our African coordinators are now familiar with big team science study demands and operations, and have perhaps grown their visibility as big team scientists in coauthoring these studies. They are more confident in using the Open Science Framework and working with the CREP platform which may inspire more CREP projects in Africa. Our African collaborators are better equipped to prepare study materials such as selecting a study for a replication study, reporting a preregistered and Registered Report study, translating and culturally adapting the measures, managing ethics review, and conducting online data collection. Nevertheless, to lead a big team science project would require consistent participation, improved capacity and resources, and a wider network and multiple consultations. There is a need for continuous capacity building for Africans working on big team science to breed the future researchers and keep the active ones productive.

2. Need to improve African research accessibility and visibility

Africans publish less than 1% in mainstream psychological science journals (Thalmayer et al., 2021). Some of the causes for the productivity gap between African and their European/North American counterparts are limited accessibility and visibility of research works. On one hand, European/North American researchers may have limited access to African literature due to 1) poor visibility of African journals, 2) most African researchers' works are not published in mainstream journals, and 3) English is the dominant language in science (Amano et al., 2023), when in fact about 30 African countries mostly communicate their research in Arabic or French. On the other hand, African researchers' access to international

journals are impacted by 1) journal subscriptions costs, e.g., research works are usually paywalled and unaffordable for African researchers and institutions, and 2) publishing in high-quality popular journals is costly. For instance, article processing charges range between \$200 and \$9000 (see Elsevier APCs price list https://legacyfileshare.elsevier.com/promis_misc/article-publishing-charge.pdf) to publish in popular scientific journals, even for hybrid and open access journals, more than most African researchers' annual salaries working in economies where the minimum wage could be as low as \$0.71 to \$248.35 per month (Katsapaos, 2023). African researchers can take advantage of the open access revolution to improve their visibility and access to their research via preprint servers and diamond open access journals, which publish and host research for free. Researchers interested in African studies can search journals and repositories (e.g., the African Journals Online, <https://www.ajol.info/index.php/ajol>; and African Digital Research Repositories, <https://www.internationalafricaninstitute.org/repositories>). Most African studies we found when searching for African claims for the ManyLabs Africa replication used qualitative methods such as focus groups and interviews, hence interested researchers may find it easier focusing on qualitative research.

3. Need to engage Africans for improved participation

Despite this dissertation focusing primarily on African thoughts and samples, African samples and researchers were underrepresented in both the main studies of the dissertation. Collaborators from only five out of fifty-four total African countries participated in the CREP Africa study. Likewise, in the ManyLabs Africa study, collaborators from just five African countries participated, contributing 1136 (18.9%) out of the total 6009 participants, and constituting 35 (27.6%) out of the 127 contributing authors. While this is a significant

improvement in African participation compared to the less than 1% samples and authors residing in Africa that is typical of mainstream psychology (Thalmayer et al., 2021), and compared to the 12.5% of participants and 2.7% of authors from the large, big-team COVID-Rapid studies (Buchanan et al., 2023; Dorison et al., 2022; Psychological Science Accelerator Self-Determination Theory Collaboration, 2022; Wang et al., 2021) even the present project could improve in including Africans, despite the fact that the inclusion of Africans was its central focus.

4. Need for more infrastructure and resources for improved productivity

A typical large-scale online replication study expenses may include participant compensation, internet access subscription, research assistants, ethics review fees, material translation, and instruments/tools. In our infrastructure and resource assessment when preparing for the CREP Africa study, we observed that only a few African collaborators and labs have some of the resources and facilities required for the studies (see the full report at <https://osf.io/gds7b>) which means we could not work with some labs. Take for instance our ethics review cost experience. While we and our African collaborators paid review fees of \$15, \$50, and \$100 for Nigerian, Kenyan, and Malawian ethics reviews, respectively, we could not afford a Tanzanian and a South African review that cost \$350 and \$917 respectively. Although a few African ethics reviews, such as in Egypt, do not require an application fee, it is ironic for African countries where researchers are poorly funded yet burdened with ethics review costs. We recommend that African research-regulating institutions such as review boards can refer to most European and North American institutions where ethics reviews are at no cost.

There is also a need for better infrastructure and facilities that causes lack of access to cutting-edge research tools and poor working conditions. In addition to lack of facilities, some

of our African collaborators struggled with heavy workloads (see also Naidoo-Chetty & du Plessis, 2021). We observed it takes relatively longer for African researchers to complete some of their tasks, such as ethics application, data collection, translation, compared to their European and North American counterparts. For instance, while most sites from the United States and Europe met their data collection target of 80 participants within two weeks, it took many African sites more than two months. We think this partly due to the poor infrastructure and tools available to our African collaborators.

Despite these realities, big team science as it is usually practiced implicitly imposes the costs and work of data collection on African research labs when they have little or no funding available and limited work time to spare. As researchers, we can do little about funding and providing facilities. However, understanding these realities helps understand how to collaborate with African researchers. Big team scientists working with the African population should communicate well-specified and feasible tasks/contributions considering the resources available to African labs/researchers and those from wealthier institutions can assist with research tools or access to small grants to aid African researchers' contribution.

5. Need to adopt open science tools and practices for credible research

While the Rottman and Young (2019) claim replicated among Africans, we failed to replicate 14 out of the 18 focal effects for the ManyLabs Africa study. Nowadays, a failed replication is unsurprising as a database of some replications showed that a number of psychology findings failed to replicate (Röseler et al., 2024) There are several open science tools and infrastructures available some of which we used for our replications and mentioned in Chapter 5. For example preregistration or Registered Report can reduce *p*-hacking, increase rigor and allow researchers to be more transparent in the conduct of these studies. We

recommend that African researchers familiarize themselves and adopt these credibility-enhancing tools and practices to produce more credible, reproducible, and replicable claims.

5. General Conclusion: Contributions

This thesis project aims to reduce the asymmetrical research participation between African and non-African populations in psychology by exploring open science practices and initiatives to improve credibility, generalizability, and inclusion in psychology. Specifically, the objectives are to 1) improve African collaborators' capacity in open science practices and big team science, 2) improve generalizability of psychology by testing the applicability and heterogeneity of psychology claims within and beyond Africa, and 3) recommend interventions to facilitate open science adoption and African participation in psychology. We were able to achieve these goals to a large extent.

In Chapter 1, our first task was to articulate the problem of underrepresentation, the generalizability problem and the big team science efforts to make psychology more generalizable. We argue that although popular big team science conducts cross-cultural investigations collectively on a large scale to expand psychology theory, a feat that individuals or a small group of researchers might not have the resources to do, these investigations are Western-focused, undersampled Africans and rarely engaged African researchers to meaningfully include Africans. On the part of Africans, big team science is nascent to African research and even if they participate in these investigations, they are hampered by the realities of limited capacity and resources. While we advocate for global psychology to better understand and nurture African psychology, our immediate approach to address to this problem was accomplished via two replications, 1) of which is a pedagogical-driven replication to familiarize African to some open science practices and infrastructures to build capacity (in Chapter 2), and 2) an African-prioritized replication to test generalizability of claims originally

discovered in Africa to non-Africans (in Chapters 3 and 4). Finally, the remaining parts of the thesis were dedicated to recommending interventions consistent with our experience conducting these investigations, the development needs and initiatives in Africa, and the credibility revolution (Chapters 5 and 6).

In Chapter 2, we conducted an Africa-wide replication via an open science-driven model to familiarize our African researchers with open science practices and build their research capacity. We partnered with CREP to execute the project in two phases. This was the first Africa-wide CREP study to provide an example of replicating a US-developed theory in Africa. In the first phase (Chapter 2a), we first created CREP syllabi and training videos on conducting a CREP study from start to finish, and organized workshops to familiarize 23 African collaborators to CREP and open science. Also, we conducted a variety of preparatory work:

- We conducted a 3-step study selection of study search, evaluation and assessment following our preset criteria of reproducibility (i.e., pre-registered and reported p -value $< .01$), ethics acceptability, applicability and relevance, and feasibility (i.e., not too complex of a study design). After which, we selected the Rottman's and Young's (2019) Studies 2 and 3 on moral judgment for the replication.
- We conducted a reanalysis of the datasets from Rottman's and Young's studies. We were able to successfully reproduce the analyses as they reported in their paper. However, Rottman's and Young's models make what we believe are implausible assumptions that produce inflated nominal rates of false positives (Judd et al., 2012). Specifically, the models omit by-person random slopes for domain, dose, and their interaction, as well as by-violation random slopes for dose. When we include the random by-person slopes for dose and the domain by dose interaction, the p -values for the focal domain by dose interaction are much larger and, in the case of Study 3, no

longer significant. However, our meta-analysis suggests that the evidence for the focal domain by dose interaction across all studies is robust, even though the evidence within studies is not quite as robust as originally reported (see Table 2).

- We ran a simulated power analysis to determine the sample size required per site. We determined that we required 10 sites to sample a minimum of 100 participants.
- We culturally adapted and translated the measures into Arabic, Chichewa, French, Portuguese, and Swahili.
- African collaborators curated individual CREP pages for their respective sites. 12 contributing sites' CREP pages were reviewed and certified according to the CREP rules. African sites provided information about their studies and procedure videos for data collection.
- We obtained an *in principle acceptance* for our Registered Replication Report after undergoing peer reviews from *Advances in Methods and Practices for the Psychological Sciences*. We then preregistered the study before we proceeded to collect data.

In the second phase (Chapter 2b), we tested the generalizability of the Rottman & Young moral transgression claims originally investigated in the United States among some Africans. Rottman and Young (2019) proposed the mere-trace theory that moral wrongness of acts that are harmful rather than impure should be more sensitive to dosage. This claim strongly replicated among our African samples meaning the Rottman's and Young's claims on moral transgression judgment were generalizable to Africans. By conducting the CREP Africa study, we were able to 1) identify the barriers and pitfalls for conducting replication studies in Africa, 2) familiarize and equip our Africans to work with the CREP model, and 3) expand and test the generalizability of the theory to Africans. However, the Africa-wide study is typical of Global North to South generalization which often leads to biases in questions being

investigated and misplaced priority for Africans. Hence, our second replication study was a preregistered replication of 3 claims originally discovered in Africa among Africans, Europeans, and North Americans. This investigation was reported in Chapters 3 and 4.

In Chapter 3a, we first moved to determine 1) studies of priority to Africans and 2) an approach for selecting these African-preferred studies for the replication. To do so, we consulted researchers working with a Global South population to determine a study focus and then developed a procedure for our study selection. Our target studies were previously-investigated social applied issues. We developed a 3-step bottom-up study selection approach of study nomination, evaluation and assessment that relied heavily on local researchers' inputs to select studies for the replication. This approach helped to identify African priorities and provided insights into their research interests (e.g., African psychology tends to focus on social applied research to proffer practical solutions to pressing issues), and methods and designs (e.g., the prevalence of qualitative methods in East African psychology).

In Chapter 3b, we selected these studies adopting our bottom-up study selection procedure as well as being guided by our predetermined criteria categorized broadly into: locality (e.g., if the study is of interest and relevant to Africans), feasibility for an online study (e.g., use instruments with few items or not a complex design), and reproducibility (e.g., well documented, reported findings with $p < .01$). In Step 1, we surveyed African researchers to nominate research areas/topics (that was also used for our search for African studies) and actual studies of interest. In Step 2, we evaluated studies nominated by African researchers and from our search using the aforementioned criteria, and shortlisted 5 studies. In Step 3, we surveyed researchers from the target study populations to assess the shortlisted studies. We observed a dichotomy of preferred studies between Africans assessors on one side and European and North American assessors on the other. However, as our replication is focused on African-prioritized studies, African opinions weighed strongly on our final study selection.

In Chapter 4, we engaged African collaborators (for the second time) as well as European and North American collaborators on pre-data collection tasks. Generally, they contributed to ethics applications to obtain approvals/authorizations, instrument cultural adaptation and translation (for non-English speaking populations), and survey/questionnaire review. Contributing sites collected data from their locals to investigate the replicability and generalizability of three sexual behavior claims on attitude towards teenage pregnancy, attitude towards condom use and forgiving extramarital sexual relationships among Africans, Europeans and North Americans. We tested the hypotheses following our preregistered methods, procedures, and analysis plans. Only 4 out of the combined 13 effects replicated and were generalizable to Africans. Region comparisons showed that Africans differ from their European and North America on these variables. We conclude that the poor replication rate and variability between African and European and North America could be due to 1) sexual liberalism or orientation difference between African and European and North American cultures and 2) questionable research practices are common in African psychology as it is presently discussed in Western psychological science. As per capacity building, such as the CREP study, the Manylabs Africa study further builds our African collaborators' capacity in open science and big team science.

In Chapter 5, we discussed the need for research development in Africa and the credibility revolution in the Western research. We recognized the research infrastructure and resources deficit in Africa and the ongoing government and development community efforts to improve the situation. Hence our discussion was to identify open infrastructure and initiatives that can promote research accessibility, infrastructure, credibility, and community, as well as help achieve the human development goals in Africa. Also, we recommend resources, training, research tools, and ethics interventions in Africa.

In Chapter 6, we provide takeaways from both the CREP Africa and Manylabs Africa studies. These takeaways are articulated in terms of the needs to better include Africans in big team science. In addition, we discuss this thesis contributions to theory, methods, capacity building, and policy below.

5.1 Theoretical contributions

The concept of replication (crisis) rate has been hugely debated in psychology in the last decade (Vazire, 2018). Beyond the replication crisis, our concern includes the underrepresentation of Africans, as well as the generalizability and applicability of psychological science findings and theories to Africans. By conducting the CREP Africa and ManyLabs Africa studies we have deepened the literature on underrepresentation of Africans and dominance of the Global North in psychology as well as the credibility revolution discourse as this may be concerned with African psychology. Overall, the replications contributed to moral judgment and sexual behavior theories' and findings' heterogeneity across Africans, Europeans and North Americans. Specifically, we found support for the Rottman and Young (2019) mere-trace hypothesis that moral judgment of purity violations are less sensitive to the severity or outcome of purity violations compared to moral judgment of harm transgressions. For the African-discovered claims replication, only 4 out of the 18 sexual behavior claims replicated and generalized to other Africans, Europeans, and North Americans. We found support for the Mgbokwere et al. (2015) claim that only 15.9% Nigerian parents agreed that parents should abort their teenage daughter pregnancy, the Teye-Kwadjo et al. (2018) claim that Ghanaian male adolescent college students reported higher rate of self-reported condom-protected sexual behavior than their female counterparts, and the Vera Cruz (2018) claims that Mozambican women were more willingness to forgive a husband involved in a less emotionally involved extramarital affair than a more involved affair, and when a cheating husband

expressed more regret than a husband that expressed less regret. On the whole, these findings underscore the need for more replication for stronger and more generalizable claims, and to understand variations across cultures, as well as for research practice corrections for better replicability, credibility and generalizability of psychological science. Lastly, we argue that it is crucial to focus on Africans for expanded and robust theory development and noted that there are infrastructural barriers that hamper Africans participation and inclusion.

5.2 Methodological contributions

This thesis also made methodology contributions to 1) study selection procedure, 2) translation procedure, 3) big-team science management workflow.

Study selection procedure. We provide two examples of study selection procedures. 1) a 3-step top-down approach used in selecting the CPA study that relied-heavily the lead researcher to nominate studies that were evaluated by a research-lead team and then assessed by local African researcher for final selection. The second one was a 4-step bottom-up for the MLA study that relied heavily on the African collaborators to nominate research themes and actual studies that guided the study search and evaluation after which local collaborators from the target populations assessed the shortlisted studies for final selection. For both study selection procedures the predetermined evaluation and assessment criteria were the same. Ideally, a study to be considered for replication should be reproducible (i.e., preregistered, reported a p -value of $<.01$, and well documented), feasible (i.e., feasible for online study, use instruments with few items, sampled the general population, and did not rely on an overly complex design), and applicable, relevant, ethically acceptable, and of interest to researchers and targeted communities. Both procedures are robust and allow researchers to involve locals at varying degrees and different stages. Top-down is good when researchers are familiar with

the literature and study population or focus on specific subjects. The bottom-up is helpful when the investigator has limited knowledge of the population and literature.

Translation procedure. Of note is our double-forward translation procedure. This is a 3-step translation procedure. Step 1, two translators provide two separate forward translations (Translations A and B). Step 2, one translator adjudicates between Translation A and Translation B statement by statement to decide which translation is more accurate, retains the original context and is meaningful to produce a combined version. Step 3, the fourth translator proofread the combined version to produce a final version. This procedure provides an alternative to the popular forward-backward translation that has been criticized for little added value for the backward translation.

Big team replication workflow. There are different tasks required for a successful large-scale replication multi-site study. Hence, effective management and workflow is then crucial. In our replications, we improved on the Psychological Science Accelerator COVID Rapid studies management sheet that includes the overall workflow of our studies. Part of workflow/tasks for CREP Africa study include a reanalysis of the original study and a simulated power analysis to evaluate the robustness of the original finding and to determine sample size respectively.²⁴ The workflow increases rigor for a replication.

5.3 Capacity building contributions

Our target is to build our African capacity in open science and big team science via replications. We contributed to materials and capacity development. For materials, we created a series for 4 CREP training videos demonstrating how to start and complete a CREP study. These training videos were subtitled into Arabic, French, Chichewa, Igbo, Swahili, and

²⁴ These tasks depend on whether the study is well documented and materials are accessible and usable. This was the case for the ManyLabs Africa study. We could not reanalyze the original study or run a power analysis for MLA replication study because we do not have adequate materials and information to complete these tasks.

Yoruba. Furthermore, materials, tools, and workflows for both studies are available on Open Science Framework for reuse under the creative commons attribution 4.0 international public license.

As par capacity, in total 57 African researchers participated in this thesis project.

- 23 African collaborators participated in our workshop on open science and the CREP model that focused on why and how to conduct a CREP study and adopting open science practices and tools.
- 38 African collaborators engaged with the Open Science Framework and CREP. They created their Open Science Framework profiles as well as forked and populated their own CREP project pages for the CREP Africa study. With this capacity and infrastructure, they are better equipped to collaborate on big team science studies, and manage and share their own research.
- 34 participated in the study selections to explored them to available approaches and workflows for selecting a study replication study,
- 26 African collaborators culturally adapted and translated the measures used in both the CREP Africa and ManyLabs Africa study into Arabic, French, Portuguese, Swahili, Chichewa, Hausa, Igbo, and Yoruba. By doing so, they further developed their translation capacity, and familiarized with our translation procedure.
- 12 African collaborators managed their local ethics review. This task enhanced their capacity in drafting, applying, reviewing, and obtaining ethics especially for an online big team science replication study.
- 25 African collaborators coordinated their local site which entailed recruiting co-investigators and leading their teams, providing required documents/materials/tools (e.g. computers, internet, ethics approval/authorization, collaboration agreement), review of survey/questionnaire, and conducting an online data collection. By doing so,

these local coordinators were able to familiarize themselves with all these contributions that are crucial for lab collaborations for big team science study.

- Our African co-authors are required to review, edit, and approve the studies' manuscripts. While some of our collaborators have been involved in manuscript writing, our studies familiarize them with writing and preparing a preregistration and Registered Replication Report manuscript.

5.4 Research policy contributions

As a matter of policy contribution, we made recommendations for policy interventions throughout this report. Broadly, research stakeholders should enable policies and provide facilities to adopt credibility-enhancing practices and conduct big-team science at organizational/managerial and individual levels. At an organizational level, management of research/teaching institutions should provide enabling infrastructure and direction for large scale studies to strive. Research-relating institutions like ethics reviews should expand their capacities and workflows to handle online large-scale (replication) research demands and implications. Scientific societies, publishers, and funders should review their policy and legal frameworks to create working conditions and develop infrastructures to enable researchers to adopt these practices. Africa capacity building initiatives can integrate the CREP Africa and ManyLabs Africa workflows into their existing training modules. At the individual level, research psychologists should adopt credibility enhancing practices and tools such as preregistration. Beyond data collection that is typical of African labs contribution, collaborators should support African collaborators to contribute measure development and cultural adaptation and translation (when possible, we recommend paying for translations), prepare materials such as data analysis script, ethics, and also support Africans to lead big team science studies. In addition, African researchers can join big team science initiatives, such as the

Psychological Science Accelerators (<http://psysciacc.org/>), a global distributed network laboratories across over 70 countries, and ManyBabies (<https://manybabies.github.io/>) a global research network focused on infant studies. African professors should furthermore review research methods curricula to include courses and subjects on open science practice and credibility-enhancing tools. All research stakeholders need to contribute to psychology to include more Africans in order to strengthen its evidence and human impact.

6. The future of African big team science

Via two big team science replications, this current thesis addressed 1) the capacity building need in open science and big team science, building much-needed expertise and networking on the continent, and 2) the underrepresentation of Africans samples and thoughts in psychology, allowing generalization of African psychology across and beyond Africa by uncovering the unknown unknowns, from a Western perspective, that may exist in African psychology. No doubt the big team science approach is a game changer in addressing generalizability and improving the credibility of psychological science. Beyond this thesis and moving forward, the future of big team science could quickly hang in the balance without the needed infrastructure, resources, and policy support from research stakeholders. Big team science is resource and expertise-intensive making it difficult to incentivize researchers to follow this path, especially for researchers in Africa where big team science is nascent, the infrastructure and resources are inadequate, and no structural change to reward big team scientists' enormous efforts. Also from a methodological perspective, big team science may quickly observe that most investigations are overpowered and need to allocate resources strategically. We think that big team scientists would be better served if they focused less on larger samples and more on sample diversity. This can be achieved if the big team approach is

widespread globally, especially to the African continent with its huge and growing population that is full of talent, potential, and latent insights.

References

- Adeniji-Neill, D. (2011). Omoluabi: The way of human being: An African philosophy's impact on Nigerian voluntary immigrants' educational and other life aspirations. *Irinkerindo: A Journal of African Migration* 5, 1-28.
- Adetula, A., Forscher, P. S., Basnight-Brown, D., Azouaghe, S., & IJzerman, H. (2022). Psychology should generalize from — not just to — Africa. *Nature Reviews Psychology*. <https://doi.org/10.1038/s44159-022-00070-y>
- Adetula, A., Forscher, P. S., Basnight-Brown, D., Azouaghe, S., Ouherrou, N., Charyate, A., Hansen, N., Adetula, G. A., & IJzerman, H. (2021). *Synergy Between the Credibility Revolution and Human Development in Africa* [Preprint]. AfricArXiv. <https://doi.org/10.31730/osf.io/e57bq>
- Adetula, A., Forscher, P. S., Basnight-Brown, D., Wagge, J. R., Namalima, T. R., Kaphesi, F. E., Kaliyapa, W., Mulungu, K., Silungwa, W., Polycarp, C., Malingumu, W., Azouaghe, S., Alsayed, E., Ndukaihe, I., Kalongonda, M., Dennis, U. D., Dzuka, A., Charyate, A., Adetula, G. A., ... IJzerman, H. (2021, July 27). *The evaluation of harm and purity transgressions in Africans: A paradigmatic replication of Rottman and Young (2019)* [Preprint]. AfricArXiv. <https://doi.org/10.31730/osf.io/hxjbu>
- Agu, I. C., Mbachu, C. O., Ezenwaka, U., Eze, I., Ezumah, N., & Onwujekwe, O. (2022). Gender norms and ideologies about adolescent sexuality: A mixed-method study of adolescents in communities, south-eastern, Nigeria. *Frontiers in Sociology*, 7. <https://doi.org/10.3389/fsoc.2022.810411>
- Ajayi, A. I., Athero, S., Muga, W., & Kabiru, C. W. (2023). Lived experiences of pregnant and parenting adolescents in Africa: A scoping review. *Reproductive Health*, 20. <https://doi.org/10.1186/s12978-023-01654-4>
- Ahlberg, B. M. (1994). Is there a distinct African sexuality? A critical response to Caldwell. *Africa: Journal of the International African Institute*, 64(2), 220–242. <https://doi.org/10.2307/1160981>

- Amano, T., Ramírez-Castañeda, V., Berdejo-Espinola, V., Borokini, I., Chowdhury, S., Golivets, M., González-Trujillo, J. D., Montañó-Centellas, F., Paudel, K., White, R. L., & Veríssimo, D. (2023). The manifold costs of being a non-native English speaker in science. *PLoS biology*, *21*(7). <https://doi.org/10.1371/journal.pbio.3002184>
- Arnett J. J. (2009). The neglected 95%: why American psychology needs to become less American. *American Psychologist*, *63*(7), 602–614. <https://doi.org/10.1037/a0016723>
- Atickem, A., Stenseth, N. Chr., Fashing, P. J., Nguyen, N., Chapman, C. A., Bekele, A., Mekonnen, A., Omeja, P. A., & Kalbitzer, U. (2019). Build science in Africa. *Nature*, *570*(7761), 297–300. <https://doi.org/10.1038/d41586-019-01885-1>
- Azouaghe, S., Adetula, A., Forscher, P. S., Basnight-Brown, D., Ouherrou, N., Charyate, A., & IJzerman, H. (2021). Psychologie et science ouverte en Afrique : pourquoi est-elle nécessaire et comment peut-on l’implanter ?. *The Journal of Quality in Education*, *11*(18), 177–193. <https://doi.org/10.37870/joqie.v11i18.274>
- Bago, B., Kovacs, M., Protzko, J., Nagy, T., Kekecs, Z., Palfi, B., Adamkovic, M., Adamus, S., Albaloooshi, S., Albayrak-Aydemir, N., Alfian, I. N., Alper, S., Alvarez-Solas, S., Alves, S. G., Amaya, S., Andresen, P. K., Anjum, G., Ansari, D., Arriaga, P., Aruta, J. J. B. R., ... Aczel, B. (2022). Situational factors shape moral judgements in the trolley dilemma in Eastern, Southern and Western countries in a culturally diverse sample. *Nature Human Behaviour*, *6*(6), 880–895. <https://doi.org/10.1038/s41562-022-01319-5>
- Bałanda-Bałdyga, A., Pilewska-Kozak, A. B., Łepecka-Klusek, C., Stadnicka, G., & Dobrowolska, B. (2020). Attitudes of teenage mothers towards pregnancy and childbirth. *International Journal of Environmental Research and Public Health*, *17*(4), 1411. <https://doi.org/10.3390/ijerph17041411>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Basen-Engquist, K., Mâsse, L. C., Coyle, K., Kirby, D., Parcel, G. S., Banspach, S., & Nodora, J. (1999). Validity of scales measuring the psychosocial determinants of HIV/STD-

- related risk behavior in adolescents. *Health Education Research*, 14(1), 25–38. <https://doi.org/10.1093/her/14.1.25>
- Bates, D. M., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100(3), 407–425. <https://doi.org/10.1037/a0021524>
- Bevan-Dye, A. L., & Akpojivi, U. (2016). South African generation Y students' self-disclosure on Facebook. *South African Journal of Psychology*, 46(1), 114–129. <https://doi.org/10.1177/0081246315602645>
- Bloom, D. E., Canning, D., Chan, K., & Luca, D. L. (2014). Higher Education and Economic Growth in Africa. *International Journal of African Higher Education*, 1(1). <https://doi.org/10.6017/ijahe.v1i1.5643>
- Boshoff, N. (2009). Neo-colonialism and research collaboration in Central Africa. *Scientometrics*, 81, 413–434. <https://doi.org/10.1007/s11192-008-2211-8>
- Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., van 't Veer, A. (2014). The replication recipe: What makes for a convincing replication?. *Journal of Experimental Social Psychology*, 50, 217–224. <https://doi.org/10.1016/j.jesp.2013.10.005>
- Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3), 389–411. <https://doi.org/10.1037/met0000159>
- Brown, J. (2010). *An introduction to overlay journals*. Repositories Support Project. <https://discovery.ucl.ac.uk/id/eprint/19081>

- Bryant, K. (2014). Research From the Global South: The important role of context in international research activities. *McGill Journal of Education*, 49, 227-235. <https://doi.org/10.7202/1025779ar>.
- Buchanan, E. M., Lewis, S., Paris, B., Forscher, P. S., Pavlacic, J. M., Beshears, J., Drexler, S. M., Gourdon-Kanhukamwe, A., Mallik, P. R., Silan, M. A. A., Miller, J. k., IJerman, H., Beaudry, J. L., Suchow, J. W., Chartier, C. R. Coles, N. A., Todsén, A. L., Levitan, C. A., Azevedo, F.,...Primbs, M. (2023). PSACR: The Psychological Science Accelerator's COVID-19 rapid-response dataset. *Scientific Data*, 10(1). <https://doi.org/10.1038/s41597-022-01811-7>
- Byers-Heinlein, K., Bergmann, C., Davies, C., Frank, M. C., Hamlin, J. K., Kline, M., Kominsky, J. F., Kosie, J. E., Lew-Williams, C., Liu, L., Mastroberardino, M., Singh, L., Waddell, C. P. G., Zettersten, M., & Soderstrom, M. (2020). Building a collaborative psychological science: Lessons learned from ManyBabies 1. *Canadian Psychology/Psychologie Canadienne*, 61(4), 349–363. <https://doi.org/10.1037/cap0000216>
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T-H., Huber, J. Johannesson, M. Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T,...Wu, H. (2018). Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nature Human Behaviour*, 2, 637-644. <https://doi.org/10.1038/s41562-018-0399-z>
- Carmack, C. C., & Lewis-Moss, R. K. (2009). Examining the theory of planned behavior applied to condom use: the effect-indicator vs. causal-indicator models. *The Journal of Primary Prevention*, 30(6), 659–676. <https://doi.org/10.1007/s10935-009-0199-3>
- Chakroff, A., Dungan, J., & Young, L. (2013). Correction: Harming ourselves and defiling others: What determines a moral domain? *PLoS ONE*, 8(9). <https://doi.org/10.1371/annotation/38818ce6-1b40-4965-aa64-b7943d2711ed>
- Cheon, B. K., Melani, I., & Hong, Y. (2020). How USA-centric is psychology? An archival study of implicit assumptions of generalizability of findings to human nature based on

origins of study samples. *Social Psychological and Personality Science*, 11(7), 928–937. <https://doi.org/10.1177/1948550620927269>

Coles, N. A., March, D. S., Marmolejo-Ramos, F., Larsen, J. T., Arinze, N. C., Ndukaihe, I. L. G., Willis, M. L., Foroni, F., Reggev, N., Mokady, A., Forscher, P. S., Hunter, J. F., Kaminski, G., Yüvrük, E., Kapucu, A., Nagy, T., Hajdu, N., Tejada, J., Freitag, R. M. K. ...Liuzza, M. T. (2022). A multi-lab test of the facial feedback hypothesis by the Many Smiles Collaboration. *Nature Human Behaviour*, 6, 1731–1742. <https://doi.org/10.1038/s41562-022-01458-9>

Cook, S. M. C., & Cameron, S. T. (2015). Social issues of teenage pregnancy. *Obstetrics, Gynaecology & Reproductive Medicine*, 25(9), 243-248. <https://doi.org/10.1016/j.ogrm.2015.06.001>

Council of Europe (2020). *Common European framework of reference for languages: learning, teaching, assessment – companion volume*. Council of Europe Publishing, Strasbourg, available at www.coe.int/lang-cefr.

Dannecker, P. (2022). Collaboration in a ‘North–South’ context: The role of power relations and the various context-based conditions. *European Journal of Development Research*, 34, 1716–1726. <https://doi.org/10.1057/s41287-022-00550-0>

Dorison, C. A., Lerner, J. S., Heller, B. H. F., Rothman, A. J., Kawachi, I. I., Wang, K., Rees, V. W., Gill, B. P., Gibbs, N., Ebersole, C. R., Tajchman, Z. V. Z., Zsido, A. N. Zrimsek, M., Chen, Z., Ziano, I., Gialitaki, Z., Ceary, C. D., Lin, Y., Kunisato, Y.,...Coles, N. A. (2022). In COVID-19 health messaging, loss framing increases anxiety with little-to-no concomitant benefits: experimental evidence from 84 countries. *Affective Science*, 3, 577–602. <https://doi.org/10.1007/s42761-022-00128-3>

Dutra, N. B. (2021). Commentary on Apicella, Norenzayan, and Henrich (2020): Who is going to run the global laboratory of the future? *Evolution and Human Behavior*, 42(3), 271-273. <https://doi.org/10.1016/j.evolhumbehav.2021.04.003>

Ebersole, C. R., Axt, J. R., & Nosek, B. A. (2016). Scientists’ reputations are based on getting it right, not being right. *PLOS Biology*, 14(5). <https://doi.org/10.1371/journal.pbio.1002460>

- Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B. V., Boucher, L., Brown, E. R., Budiman, N. I., Cairo, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., . . . Nosek, B. A. (2016). Many labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology, 67*, 68–82. <https://doi.org/10.1016/j.jesp.2015.10.012>
- Emery, L. F., Silverman, D. M., Carey, R. M. (2023). Conducting research with people in lower-socioeconomic-status contexts. *Advances in Methods and Practices in Psychological Science, 6*(4). <https://doi.org/10.1177/25152459231193044>
- Eze, M. O. (2010). Ubuntu: Toward a new public discourse. In M. O. Eze (Eds.) *Intellectual History in Contemporary South Africa*, (pp. 181–192). Palgrave Macmillan.
- Fearon, J. D. (2003). Ethnic and Cultural Diversity by Country. *Journal of Economic Growth, 8*, 195–222.
- Fecher, B., & Friesike, S. (2014). Open Science: One Term, Five Schools of Thought. In S. Bartling & S. Friesike (Eds.), *Opening Science* (pp. 17–47). Springer International Publishing. https://doi.org/10.1007/978-3-319-00026-8_2
- Flanagan, E. A., Erickson, H. C., Parchem, S. J., Smith, C. V., Poland, N., Nelson, S. C., & Redman, S. D. (2021). Gender differences in attitudes and behaviors toward condoms and birth control among Chicago adolescents. *World Medical & Health Policy, 13*, 675–694. <https://doi.org/10.1002/wmh3.422>
- Forscher, P. S., Basnight-Brown, D. M., Dutra, N., Adetula, A., Silan, M., & IJzerman, H. (2021, December 29). Psychological Science Needs the Entire Globe, Part 3: Let’s Does Psychological Science Deserve Brilliant Researchers From Outside North America and Europe? *APS Observer 35*(1). <https://www.psychologicalscience.org/observer/entire-globe-part-3>
- Forscher, P. S., Wagenmakers, E.-J., Coles, N. A., Silan, M. A., Dutra, N., Basnight-Brown, D., & IJzerman, H. (2022). The benefits, barriers, and risks of Big-Team Science. *Perspectives on Psychological Science, 18*(3), 607–623. <https://doi.org/10.1177/17456916221082970>

- Francis, G. (2012). Too good to be true: Publication bias in two prominent studies from experimental psychology. *Psychonomic Bulletin & Review*, *19*(2), 151–156. <https://doi.org/10.3758/s13423-012-0227-9>
- Galal, S. (2024, March 14). Internet penetration in Africa January 2024, by country. *Statista*. <https://www.statista.com/statistics/1124283/internet-penetration-in-africa-by-country/>
- Gervais, W. M. (2021). Practical methodological reform needs good theory. *Perspectives on Psychological Science*, *16*(4), 827–843. <https://doi.org/10.1177/1745691620977471>
- Ghai, S., Forscher, P. S., & Hu, C-P. (2023, May 16). Big-team science does not guarantee generalizability [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/avcsp>
- Giner-Sorolla, R., & Chapman, H. A. (2017). Beyond Purity: Moral Disgust Toward Bad Character. *Psychological Science*, *28*(1), 80–91. <https://doi.org/10.1177/0956797616673193>
- Green, D. (2021, September 8). North-South power differentials and competition in the research business. *From Poverty to Power*. <https://frompoverty.oxfam.org.uk/north-south-power-differentials-and-competition-in-the-research-business/>
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*(5827), 998–1002. <https://doi.org/10.1126/science.1137651>
- Hallsworth, M. (2023). A manifesto for applying behavioural science. *Nature Human Behaviour*, *7*, 310–322. <https://doi.org/10.1038/s41562-023-01555-3>
- Hearty, M. (2022, August 22). Kenya Launches Agenda to Harmonize Research in Reproductive Health. *Science Africa*. <https://scienceafrica.co.ke/2022/08/04/kenya-launches-agenda-to-harmonize-research-in-reproductive-health/>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. <https://doi.org/10.1017>

- Hofstede, G., Hofstede, G.J. and Minkov, M. (2010). *Cultures and Organizations: Software of the Mind: Intercultural Cooperation and Its Importance for Survival* (2nd ed.), McGraw-Hill, London.
- Holcombe, A. O. (2019). Contributorship, not authorship: Use CRediT to indicate who did what. *Publications*, 7(3), 48. <https://doi.org/10.3390/publications7030048>
- Holland, K. J., & French, S. E. (2012). Condom negotiation strategy use and effectiveness among college students. *Journal of Sex Research*, 49(5), 443–453. <https://doi.org/10.1080/00224499.2011.568128>
- Hone, L. S. E., Testa, M., & Wang, W. (2020). Sociosexuality and sex with new partners: Indirect effects via drinking at parties and bars. *Journal of Studies on Alcohol and Drugs*, 81(5), 624–630. <https://doi.org/10.15288/jsad.2020.81.624>
- Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, 97(6), 963–976. <https://doi.org/10.1037/a0017423>
- Hountondji, P. (1990). Scientific dependence in Africa today. *Research in African Literatures*, 21(3), 5–15.
- Huang, C.-K. (Karl), Neylon, C., Hosking, R., Montgomery, L., Wilson, K. S., Ozaygen, A., & Brookes-Kenworthy, C. (2020). Evaluating the impact of open access policies on research institutions. *ELife*, 9, e57067. <https://doi.org/10.7554/eLife.57067>
- Ibrahim, A. S. (2011). Arab world psychology. In K. D. Keith (Eds.), *The Encyclopedia of Cross-Cultural Psychology*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118339893.wbeccp032>
- ICEF Monitor. (2015, March 16). *African summit calls for major expansion of higher education*. <https://monitor.icef.com/2015/03/african-summit-calls-for-major-expansion-of-higher-education/>
- IJzerman, H., Blanken, I., Brandt, M. J., Oerlemans, J. M., Van den Hoogenhof, M. M. W., Franken, S. J. M., & Oerlemans, M. W. G. (2014). Sex differences in distress from

- infidelity in early adulthood and in later life: A replication and meta-analysis of Shackelford et al. (2004). *Social Psychology*, 45(3), 202–208. <https://doi.org/10.1027/1864-9335/a000185>
- IJzerman, H., Dutra, N., Silan, M., Adetula, A., Basnight-Brown, D. M., & Forscher, P. S. (2021, August 30). Psychological science needs the entire Globe, Part 1: The problem with U.S. dominance in psychological science. *APS Observer* 34(5). <https://www.psychologicalscience.org/observer/global-psych-science>
- Ioannidis J. P. (2005). Why most published research findings are false. *PLoS medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Jarke, H., Anand-Vembar, S., Alzahawi, S., Andersen, T. L., Bojanić, L., Carstensen, A., Feldman, G., Garcia-Garzon, E., Kapoor, H., Lewis, S., Todsén, A. L., Večkalov, B., Zickfeld, J. H., & Geiger, S. J. (2022). A roadmap to large-scale multi-country replications in psychology. *Collabra: Psychology*, 8(1). <https://doi.org/10.1525/collabra.57538>
- Jemmott, J. B., 3rd, Heeren, G. A., Ngwane, Z., Hewitt, N., Jemmott, L. S., Shell, R., & O'leary, A. (2007). Theory of planned behaviour predictors of intention to use condoms among Xhosa adolescents in South Africa. *AIDS care*, 19(5), 677–684. <https://doi.org/10.1080/09540120601084308>
- Jimu, S. E., Ntoimo, L. F. C., & Okonofua, F. E. (2023). Prevalence and determinants of condom use among the youth in Malawi: evidence from the 2015/16 Malawi Demographic and Health Survey. *Reproductive Health* 20. <https://doi.org/10.1186/s12978-023-01714-9>
- Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., Ndukaihe, I. L. G., Bloxson, N. G., Lewis, S. C., Foroni, F., Willis, M. L., Cubillas, C. P., Vadillo, M. A., Turiegano, E., Gilead, M., Simchon, A., Saribay, S. A., Owsley, N. C., Jang, C., Mburu, G., ... Coles, N. A. (2021). To which world regions does the valence-dominance model of social perception apply?. *Nature human behaviour*, 5(1), 159–169. <https://doi.org/10.1038/s41562-020-01007-2>

- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. <https://doi.org/10.1037/a0028347>
- Kaplan, R. M., & Irvin, V. L. (2015). Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time. *PLOS One*, *10*(8), e0132382. <https://doi.org/10.1371/journal.pone.0132382>
- Katsapaos, K. (2023, May 26). Minimum wage by country: Global payroll guide. *Mauve*. <https://mauvegroup.com/innovation-hub/blog/minimum-wage-by-country-global-payroll-guide>
- Kenmoe, S. (2022). African fictions to boost the scientific culture in Africa. *Nature Reviews Materials*, *7*, 675–676. <https://doi.org/10.1038/s41578-022-00474-5>
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, *2*(3), 196-217. https://doi.org/10.1207/s15327957pspr0203_4
- Kieslinger, B., Schäfer, T., Heigl, F., Dörler, D., Richter, A., & Bonn, A. (2018). Evaluating citizen science - Towards an open framework. In: S. Hekler, M. Haklay, A. Bowser, J. Vogel & A. Bonn, (Eds.), *Citizen Science - Innovation in Open Science, Society and Policy*. (pp. 81-95). UCL Press. <https://doi.org/10.14324/111.9781787352339>
- Kijilian, C., Sharifian, M., Paris, B., Iyer, A., & Silan, M. A. (2022, September 2). *PSA Study Capacity Report 2022-2023*. <https://doi.org/10.31234/osf.io/7krhn>
- Kirya, M. (2019). Corruption in universities: *Paths to integrity in the higher education subsector*. *Anti-Corruption Resource Centre*. https://data.opendevlopmentmekong.net/dataset/0d23d14c-5ec4-45ca-95b0-d4997e7dd713/resource/d0daa8be-8cd3-44f4-9d47-be2af130a324/download/u4_corruption-in-universities_paths-to-integrity-in-the-higher-education-subsector.pdf

- Klein, R. A., Cook, C. L., Ebersole, C. R., Vitiello, C., Nosek, B. A., Hilgard, J., Ahn, P. H., Brady, A. J., Chartier, C. R., Christopherson, C. D., Clay, S., Collisson, B., Crawford, J. T., Cromar, R., Gardiner, G., Gosnell, C. L., Grahe, J., Hall, C., Howard, I., ... & Ratliff, K. A. (2022). Many Labs 4: Failure to replicate mortality salience effect with and without original author involvement. *Collabra: Psychology*, 8(1), 35271. <https://doi.org/10.1525/collabra.35271>
- Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, Š., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., Cemalcilar, Z., Chandler, J., Cheong, W., Davis, W. E., Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E. M., . . . Nosek, B. A. (2014). Investigating variation in replicability: A “many labs” replication project. *Social Psychology*, 45(3), 142–152. <https://doi.org/10.1027/1864-9335/a000178>
- Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Jr., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., ... Nosek, B. A. (2018). Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, 1(4), 443–490. <https://doi.org/10.1177/2515245918810225>
- Kok, M. O., Gyapong, J. O., Wolffers, I., Ofori-Adjei, D., & Ruitenberg, E. J. (2017). Towards fair and effective North–South collaboration: realising a programme for demand-driven and locally led research. *Health Research Policy and System*, 15. <https://doi.org/10.1186/s12961-017-0251-3>
- Kombe, F., Anunobi, E. N., Tshifugula, N. P., Wassenaar, D., Njadingwe, D., Mwalukore, S., Chinyama, J., Randrianasolo, B., Akindeh, P., Dlamini, P. S., Ramiandrisoa, F. N., & Ranaivo, N. (2014). Promoting research integrity in Africa: An African voice of concern on research misconduct and the way forward. *Developing World Bioethics*, 14(3), 158–166. <https://doi.org/10.1111/dewb.12024>
- Kombo, S. (n.d.). *Using behavioural informed communication to drive civic engagement*. [Unpublished paper] <https://work.busaracenter.org/case-studies/A11%20Community%20Messaging%20&%20Civic%20Participation%20Pha>

[se%20%20-%20Uganda.pdf?_ga=2.147651319.1196572791.1657005992-152356769.1647274662](https://doi.org/10.1038/s44271-023-00003-2)

- Korbmacher, M., Azevedo, F., Pennington, C. R., Hartmann, H., Pownall, M., Schmidt, K., Elsherif, M., Breznau, N., Robertson, O., Kalandadze, T., Yu, S., Baker, B. J., O'Mahony, A., Olsnes, J. Ø. -S., Shaw, J. J., Gjoneska, B., Yamada, Y., Röer, J. P., Murphy, J., ... Evans, T. (2023). The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology, 1*(3). <https://doi.org/10.1038/s44271-023-00003-2>
- Kumwenda, S., Niang, E. H. A., Orondo, P. W., Pote, W., Oyafajo, L. A., Bongo, G. N., & Chiwona, B. (2017). Challenges facing young African scientists in their research careers: a qualitative exploratory study. *Malawi Medical Journal: The Journal of Medical Association of Malawi, 29*, 1-4. <https://doi.org/10.4314/mmj.v29i1.1>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lakati, A. S. & Masibo, P. K. (2023). A call for African universities to define their research priorities. *The Lancet: Global Health, 11*(10), e1505-e1506. [https://doi.org/10.1016/S2214-109X\(23\)00366-2](https://doi.org/10.1016/S2214-109X(23)00366-2)
- Lakens, D. (2019). The Value of Preregistration for Psychological Science: A Conceptual Analysis. *Japanese Psychological Review, 62*(3), 221–230. https://www.jstage.jst.go.jp/article/sjpr/62/3/62_221/_pdf
- Lakens, D. (2022). Sample size justification. *Collabra: Psychology, 8*(1). 33267. <https://doi.org/10.1525/collabra.33267>
- Langham-Putrow, A., Bakker, C., & Riegelman, A. (2021). Is the open access citation advantage real? A systematic review of the citation of open access and subscription-based articles. *PloS one, 16*(6), e0253129. <https://doi.org/10.1371/journal.pone.0253129>

- Lawal, G. (2007). Corruption and development in Africa: Challenges for political and economic change. *Humanity and Social Sciences Journal*, 2, 1–7.
- Lee, J. C., Hall, D. L., & Wood, W. (2018). Experiential or material purchases? Social class determines purchase happiness. *Psychological Science*, 29(7), 1031–1039. <https://doi.org/10.1177/0956797617736386>
- Looney, P., Corzine, G., Daly, N., Chambers, S., Crosby, J., & Lathrop, J. (n.d.). *Does dosage effect evaluations of purity transgressions versus harm transgressions? A replication of Rottman and Young (2019): Specks of Dirt and Tons of Pain*. https://moodle.swarthmore.edu/pluginfile.php/616880/mod_resource/content/1/PSYC%20025%20Study%202.pdf
- Matenga, T. F. L., Zulu, J. M., Corbin, J. H., & Mweemba, O. (2019). Contemporary issues in north-south health research partnerships: perspectives of health research stakeholders in Zambia. *Health Research Policy and Systems*, 17(1), 7. <https://doi.org/10.1186/s12961-018-0409-7>
- Meskó, N. & Láng, A. (2013). Women evaluate their partners' betrayal: Emotional reactions, forgiveness, perceived threats of infidelity. In E. Farkas (Ed.), *Acta Szekszardiensium/Scientific Publications, Tom XV*. (pp. 47-59.) Szekszárd: Illyés Gyula Faculty, University of Pécs. <http://www.evulociospszichologia.hu/uploads/9/3/4/5/9345921/mesko-lang.2013.betrayal.pdf>
- Mgbokwere, D. O., Esienumoh, E. E., & Uyana, D. A. (2015). Perception and attitudes of parents towards teenage pregnancy in a rural community of Cross River state, Nigeria. *Global Journal of Pure and Applied Sciences*, 21(2), 181. <https://doi.org/10.4314/gjpas.v21i2.10>
- Michalopoulos, S., & Papaioannou, E. (2020). Historical legacies and African development. *Journal of Economic Literature*, 58(1), 53–128. <https://doi.org/10.1257/jel.20181447>
- Moshontz, H., Binion, G. E., Walton, H., Brown, B. T., & Syed, M. (2021). A guide to posting and managing preprints. *Advances in Methods and Practices in Psychological Science*, 4(2). <https://doi.org/10.1177/25152459211019948>

- Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H. Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. J., Keene, J. R., Protzko, J., Aczel, B.,...& Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science*, 1(4), 501-515. doi:[10.1177/2515245918797607](https://doi.org/10.1177/2515245918797607)
- Mouton, J., & Valentine, A. (2017). The extent of South African authored articles in predatory journals. *South African Journal of Science*, 113(7/8). <https://doi.org/10.17159/sajs.2017/20170010>
- Mpofu, E. (2002). Psychology in sub-Saharan Africa: Challenges, prospects and promises. *International Journal of Psychology*, 37(3), 179–186. <https://doi.org/10.1080/00207590244000061>
- Mughogho, W., Adhiambo, J., & Forscher, P. S. (2023). African researchers must be full participants in behavioural science research. *Nature Human Behaviour* 7, 297–299. <https://doi.org/10.1038/s41562-023-01536-6>
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 0021. <https://doi.org/10.1038/s41562-016-0021>
- Musau, Z. (2017, December). Africa grapples with huge disparities in education. *Africa Renewal*. <https://www.un.org/africarenewal/magazine/december-2017-march-2018/africa-grapples-huge-disparities-education>
- Mwangi, K. W., Mainye, N., Ouso, D. O., Esoh, K., Muraya, A. W., Mwangi, C. K., Naitore, C., Karega, P., Kibet-Rono, G., Musundi, S., Mutisya, J., Mwangi, E., Mgawe, C., Miruka, S., Kibet, C. K., & OpenScienceKE Collaborators (2021). Open Science in Kenya: Where Are We?. *Frontiers in research metrics and analytics*, 6, 669 - 675. <https://doi.org/10.3389/frma.2021.669675>

- Nabyonga-Orem, J., Asamani, J. A., Nyirenda, T., & Abimbola, S. (2020). Article processing charges are stalling the progress of African researchers: a call for urgent reforms. *BMJ Global Health*, 5(9). <https://doi.org/10.1136/bmjgh-2020-003650>
- Naidoo-Chetty, M., & du Plessis, M. (2021). Job demands and job resources of academics in higher education. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.631171>
- Nathan, S. F., Berglas, N. F., Kaller, S., Mays, A., & Biggs, M. A. (2023). Reasons for having unprotected sex among adolescents and young adults accessing reproductive health services. *Women's health issues : official publication of the Jacobs Institute of Women's Health*, 33(3), 222–227. <https://doi.org/10.1016/j.whi.2022.11.006>
- Nordling, L. (2018). World Bank pours hundreds of millions into African science. *Nature*, 561(7721), 16–16. <https://doi.org/10.1038/d41586-018-06094-w>
- Nsamenang, A. B. (2007). Origins and development of scientific psychology in Afrique Noire. In *Psychology: IUPsyS Global Resource* (2007th ed.). Psychology Press. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.522.4330&rep=rep1&type=pdf>
- Nosek, B. A., & Errington, T. M. (2020). What is replication?. *PLoS biology*, 18(3). <https://doi.org/10.1371/journal.pbio.3000691>
- Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, 73(1), 719–748. <https://doi.org/10.1146/annurev-psych-020821-114157>
- Ojedokun, I. M. (2015). Extramarital affair as correlate of reproductive health and home instability among couples in Ibadan, Nigeria. *African Journal of Social Work*, 5(2), 1–40. <https://www.ajol.info/index.php/ajsw/article/download/133938/123557>

- Ojiji, O. O. (2015). Fifty years of psychology in Nigeria: Are we still teaching science or folktales? *African Journal for the Psychological Studies of Social Issues*, 18(2). <http://www.ajpssi.org/index.php/ajpssi/article/view/129>
- Onie, S. (2020). Redesign open science for Asia, Africa and Latin America. *Nature*, 587(7832), 35–37. <https://doi.org/10.1038/d41586-020-03052-3>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), 1–8. <https://doi.org/10.1126/science.aac4716>
- Oppong, S. (2022). Indigenous psychology in Africa: Centrality of culture, misunderstandings, and global positioning. *Theory & Psychology*, 32(6), 953-973. <https://doi.org/10.1177/09593543221097334>
- Penke, L. (2010). Revised sociosexual orientation inventory. In T. D. Fisher, C. M. Davis, W. L. Yarber, & S. L. Davis (Eds.), *Handbook of sexuality-related measures* (3rd ed., pp. 622-625). Routledge.
- Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, 95(5), 1113–1135. <https://doi.org/10.1037/0022-3514.95.5.1113>
- Pillay, S. R. (2017). Cracking the fortress: can we really decolonize psychology? *South African Journal of Psychology*, 47(2), 135-140. <http://doi:10.1177/0081246317698059>
- Premachandra, B., & Lewis, N. A. (2022). Do we report the information that is necessary to give psychology away? A scoping review of the psychological intervention literature 2000–2018. *Perspectives on Psychological Science*, 17(1), 226-238. <https://doi.org/10.1177/1745691620974774>
- Psychological Science Accelerator Self-Determination Theory Collaboration (2022). A global experiment on motivating social distancing during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, 119(22). <https://doi.org/10.1073/pnas.2111091119>

- Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of Homo sapiens: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences*, *115*(45), 11401-11405. <https://doi.org/10.1073/Pnas.1721165115>
- Rohatgi, A. (2024). WebPlotDigitizer 4.7. <https://automeris.io/WebPlotDigitizer.html>
- Rohwer, A., Young, T., Wager, E., & Garner, P. (2017). Authorship, plagiarism and conflict of interest: Views and practices from low/middle-income country health researchers. *BMJ Open*, *7*(11), e018467. <https://doi.org/10.1136/bmjopen-2017-018467>
- Röseler, L., Kaiser, L., Doetsch, C. A., Klett, N., Seida, C., Schütz, A., Aczel, B., Adelina, N., Agostini, V., Alarie, S., Albayarak-Aydemir, N. Aldoh, A., H. Al-Hoorie, A. Azevedo, F., Baker, B. J., Barth, C. L., Beitner, J., Brick, C., Brohmer, H....,Zhang, Y., (2024, April 11). *The Replication Database: Documenting the Replicability of Psychological Science* [Preprint]. MetaArXiv. <https://doi.org/10.31222/osf.io/me2ub>
- Roser, M. (2013). *Economic Growth*. Our World In Data. <https://ourworldindata.org/economic-growth>
- Roser, M., Ritchie, H., & Ortiz-Ospina, E. (2013). *World Population Growth*. Our World In Data. <https://ourworldindata.org/world-population-growth>
- Rottman, J., & Young, L. (2019). Specks of dirt and tons of pain: Dosage distinguishes impurity from harm. *Psychological Science*, *30*(8), 1151–1160. <https://doi.org/10.1177/0956797619855382>
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, *76*(4), 574–586. <https://doi.org/10.1037/0022-3514.76.4.574>
- Scheel, A. M., Schijen, M., & Lakens, D. (2020). *An excess of positive results: Comparing the standard Psychology literature with Registered Reports* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/p6e9c>

- Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods*, 17(4), 551–566. <https://doi.org/10.1037/a0029487>
- Sfiligoi, I., Bradley, D. C., Holzman, B., Mhashilkar, P., Padhi, S., & Wurthwein, F. (2009). The Pilot Way to Grid Resources Using glideinWMS. *2009 WRI World Congress on Computer Science and Information Engineering*, 428–432. <https://doi.org/10.1109/CSIE.2009.950>
- Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, 69(1), 487–510. <https://doi.org/10.1146/annurev-psych-122216-011845>
- Sharra, S. (2023, November 2). Steadying the edifice: Rebuilding Malawi’s education sector. *University World News: Africa Edition*. <https://www.universityworldnews.com/post.php?story=20231101074113623>
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, 12(6), 1123–1128. <https://doi.org/10.1177/1745691617708630>
- Simpson, J. A., & Gangestad, S. W. (1991). Individual differences in sociosexuality: evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology*, 60(6), 870–883. <https://doi.org/10.1037//0022-3514.60.6.870>
- Singh, L., Killen, M., and Smetana, J. G. (2023, August 31). Global science requires greater equity, diversity, and cultural precision. *APS Observer* 36(5). <https://www.psychologicalscience.org/observer/gs-equity-diversity-cultural-precision>
- Smith, A. P. (2000). The journal as an overlay on preprint databases. *Learned Publishing*, 13(1), 43–48. <https://doi.org/10.1087/09531510050145542>
- Spellman, B. A. (2015). A short (personal) future history of revolution 2.0. *Perspectives on Psychological Science*, 10(6), 886–899. <https://doi.org/10.1177/1745691615609918>

- Stankov, L., & Lee, J. (2016). Nastiness, morality and religiosity in 33 nations. *Personality and Individual Differences, 99*, 56–66. <https://doi.org/10.1016/j.paid.2016.04.069>
- Teye-Kwadjo, E., Kagee, A., & Swart, H. (2018). Condom use negotiation among high school adolescents in Ghana: The role of gender. *South African Journal of Psychology, 48*(4), 465–475. <https://doi.org/10.1177/0081246317730282>
- Thalmayer, A. G., Toscanelli, C., & Arnett, J. J. (2021). The neglected 95% revisited: Is American psychology becoming less American?. *American Psychologist, 76*(1), 116–129. <https://doi.org/10.1037/amp0000622>
- The AREN team. (2023, August 1). African Reproducibility Network (AREN): Ambassadors bridging the gaps in open science adoption and advocacy across Africa. *eLife Sciences Publications*. <https://elifesciences.org/inside-elifesciences/820c1044/african-reproducibility-network-aren-ambassadors-bridging-the-gaps-in-open-science-adoption-and-advocacy-across-africa>
- The World Bank Group & Elsevier. (2014). *A Decade of Development in Sub-Saharan African Science, Technology, Engineering & Mathematics Research* (No. 91016). <https://doi.org/worldbank.org/africa/stemresearchreport>
- Tiokhin, L., Hackman, J., Munira, S., Jesmin, K., & Hruschka, D. (2019). Generalizability is not optional: Insights from a cross-cultural study of social discounting. *Royal Society Open Science, 6*(2), 181386. <https://doi.org/10.1098/rsos.181386>
- Ubillos, S., Páez, D., & González, J. L. (2000). Culture and sexual behavior. *Psicothema, 12*(1), 70-82. <https://www.psicothema.com/pdf/399.pdf>
- United Nations Innovation Network. (2022, July 15). *United Nations Behavioural Science Report*. <https://drive.google.com/file/d/1cyxNkglSChdGQOomR-dLm4ZNkH-Mjos/view>
- United Nations Educational, Scientific and Cultural Organization, (UNESCO; 2021). UNESCO Recommendation on Open Science. <https://doi.org/10.54677/MNMH8546>

United Nations Educational, Scientific and Cultural Organization, (UNESCO; 2021). *UNESCO Science Report: The race against time for smarter development; executive summary* (SC-2021/WS/7). <https://unesdoc.unesco.org/ark:/48223/pf0000377250>

United Nations Educational, Scientific and Cultural Organization, (UNESCO; 2023). *Priority Africa in action, 2022-2023*. <https://unesdoc.unesco.org/ark:/48223/pf0000386986>

United Nations, Department of Economic and Social Affairs, Population Division (2022). *World Population Prospects 2022*. Data Sources. UN DESA/POP/2022/DC/NO. 9.

Van der Veken, K., Belaid, L., Delvaux, T. & De Brouwere, D. (2017). Research capacity building through North–South–South networking: towards true partnership? An exploratory study of a network for scientific support in the field of sexual and reproductive health. *Health Research Policy and System*, 15. <https://doi.org/10.1186/s12961-017-0202-z>

Vazire, S. (2017). Quality uncertainty erodes trust in science. *Collabra: Psychology*, 3(1), 1. <https://doi.org/10.1525/collabra.74>

Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science*, 13(4), 411–417. <https://doi.org/10.1177/1745691617751884>

Vera Cruz, G. (2018). Would Mozambican women really “tolerate” their husbands’ extramarital sexual relationships as socially recommended?. *Sexuality & Culture*, 22, 1263–1278. <https://doi.org/10.1007/s12119-018-9520-8>

Visser, I., Bergmann, C., Byers-Heinlein, K., Dal Ben, R., Duch, W., Forbes, S., Franchin, L., Frank, M. C., Geraci, A., Hamlin, J. K., Kaldy, Z., Kulke, L., Lavery, C., Lew-Williams, C., Mateu, V., Mayor, J., Moreau, D., Nomikou, I., Schuwerk, T., Simpson, E. A., ... Zettersten, M. (2022). Improving the generalizability of infant psychological research: The ManyBabies model. *Behavioral and Brain Sciences*, 45. <https://doi.org/10.1017/S0140525X21000455>

Vohs, K. D., Schmeichel, B. J., Lohmann, S., Gronau, Q. F., Finley, A. J., Ainsworth, S. E., Alquist, J. L., Baker, M. D., Brizi, A., Bunyi, A., Butschek, G. J., Campbell, C., Capaldi,

- J., Cau, C., Chambers, H., Chatzisarantis, N. L. D., Christensen, W. J., Clay, S. L., Curtis, J., ... Albarracín, D. (2021). A multisite preregistered paradigmatic test of the ego-depletion effect. *Psychological Science*, 32(10), 1566-1581. <https://doi.org/10.1177/0956797621989733>
- Wagge, J. R., Baciú, C., Banas, K., Nadler, J. T., Schwarz, S., Weisberg, Y., IJzerman, H., Legate, N., & Grahe, J. (2019). A demonstration of the Collaborative Replication and Education Project: Replication attempts of the red-romance effect. *Collabra: Psychology*, 5(1), 5. <https://doi.org/10.1525/collabra.177>
- Wagge, J. R., Brandt, M. J., Lazarevic, L. B., Legate, N., Christopherson, C., Wiggins, B., & Grahe, J. E. (2019). Publishing research with undergraduate students via replication work: The Collaborative Replications and Education Project. *Frontiers in Psychology*, 10, 247. <https://doi.org/10.3389/fpsyg.2019.00247>
- Walsh, J. L., Senn, T. E., Scott-Sheldon, L. A., Vanable, P. A., & Carey, M. P. (2011). Predicting condom use using the Information-Motivation-Behavioral Skills (IMB) model: a multivariate latent growth curve analysis. *Annals of Behavioral Medicine : a publication of the Society of Behavioral Medicine*, 42(2), 235-244. <https://doi.org/10.1007/s12160-011-9284-y>
- Wang, K., Goldenberg, A., Dorison, C.A., Miller, J. K., Uusberg, A., Lerner, J. S., Gross, J. J., Agesin, B. B., Bernardo, M., Campos, O., Eudave, L., Grzech, K., Ozery, D. H., Jackson, E. A., Garcia, E. O. L., Drexler, S. M., Jurković, A. P., Rana, K., Wilson, J. P., ...Moshontz, H. (2021). A multi-country test of brief reappraisal interventions on emotions during the COVID-19 pandemic. *Nature Human Behaviour* 5, 1089-1110. <https://doi.org/10.1038/s41562-021-01173-x>
- Wareham, C. S. (2017). A duty to explore African ethics? *Ethical Theory and Moral Practice*, 20(4), 857-872. <https://doi.org/10.1007/s10677-017-9826-x>
- Webb, L., Kyaddondo, D., Ford, T., Bergqvist, A., & Cox, N. (2023). Psychosocial health in adolescent unmarried motherhood in rural Uganda: Implications for community-based collaborative mental health education, and empowerment strategies in the prevention

of depression and suicide. *Transcultural Psychiatry*, 60(3), 537-551.
<https://doi.org/10.1177/13634615221147361>

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, 143(5), 2020–2045.
<https://doi.org/10.1037/xge0000014>

World Health Organization (WHO; March 2023). *Health researchers (in full-time equivalent) per million inhabitants, by WHO Region (first set of charts)*.
<https://www.who.int/observatories/global-observatory-on-health-research-and-development/indicators/health-researchers-in-full-time-equivalent-per-million-inhabitants-by-who-region-first-set-of-charts>

World Health Organization (2022, September 30). Expanding capacity in behavioural insights in the African Region. <https://www.who.int/news/item/30-09-2022-expanding-capacity-to-apply-behavioural-science-in-the-african-region>

Xia, J., Harmon, J. L., Connolly, K. G., Donnelly, R. M., Anderson, M. R., & Howard, H. A. (2015). Who publishes in “predatory” journals? *Journal of the Association for Information Science and Technology*, 66(7), 1406–1417.
<https://doi.org/10.1002/asi.23265>

Young, C. (2012). *The postcolonial state in Africa: Fifty years of independence, 1960-2010*. The University of Wisconsin Press.

Zaadnoordijk, L., Buckler, H., Cusack, R., Tsuji, S., & Bergmann, C. (2021). A global perspective on testing infants online: Introducing ManyBabies-AtHome. *Frontiers in Psychology*, 12: 703234. <http://doi.org/10.3389/fpsyg.2021.703234>

Zeineddine, F. B., Saab, R., Láštiová, B. Kende, A., & Ayanian, A. H. (2022). “Some uninteresting data from a faraway country”: Inequity and coloniality in international social psychological publications. *Journal of Social Issues*, 78, 320– 345.
<https://doi.org/10.1111/josi.12481>

Appendix I

1. Consultants' general opinions of the proposed ManyLabs Africa study

Generally, consultants opined that the proposed ManyLabs Africa study is exciting and can help advance research practices and theory in the global south. However, the replication should focus on social issue claims with practical application in Africa. Below is a point-by-point summary of the consultants, anonymized and randomly ordered from topic-to-topic.

- Research in developing countries should focus on “national issues” such as poverty, food security, corruption, et cetera. However, these “national issues” faced by developing countries are rarely of concern to researchers or the public in developed countries. Two critical issues of concern are 1) determining a common/shared study interest for developing countries (Africa) and developed countries (Europe and North America) populations, and 2) deciding what research designs are appropriate for a “shared research subject” replication for these populations.
- The proposed study should focus on real-life and not convenient studies, such as conducting field experiment research instead of an online survey study. However, we noted during our discussion that field experiment studies could be time-consuming. Perhaps limiting the ManyLabs Africa study to African populations only will ensure feasible field experimental studies.
- The ManyLabs Africa proposed an interesting approach to studying the African populations. However, Consultant II questioned adopting a single approach for studying hugely diverse African populations with the existing disparities in African research cultures. However, some North African countries such as Tunisia, Morocco, and Algeria have research interests focused heavily on applied research. Also, clinical

psychology dominates most African countries' psychology research interests and practices.

- Another issue raised concerns the application value of these studies. The consultant opined that determining the applicability of these studies is crucial to getting the best-fit study for the study populations and collaborators.
- In line with which research area and what applies to Africans, the consultant described behavioral health issues as an area of research that would interest African researchers and communities.
- However, the consultant cautioned that selecting a study to replicate via African journals might be challenging. There are few high-quality journals available in Africa due to poor resources and infrastructure and lack of rigor, and some of these African journals engage in questionable practices.
- Also, there is a need to consider unethical research practices and methodology issues.
- The project could be very important for theorizing and help provide answers to understanding the universality of human behavior. Also, this will be an interesting project because there are few crowdsourced multisite research projects in Africa. However, noted some challenges/barriers working with African researchers:
 - i. African researchers' workload is heavy, with extended hours of teaching and administrative roles that stress their scarce resources—making it more demanding for African collaborators to participate in these studies.
 - ii. Getting samples can be problematic. Sometimes collaborators require research assistants to be able to meet up with the required number of samples. Also, the easy-to-reach population for study samples in Africa is college students. Lastly, incentivizing participants can help encourage participation.

iii. Unlike in Europe, online resources and facilities are not readily available in most developing countries in Asia, Latin America, and even some African countries like Uganda, Kenya, and Malawi, which means there are limitations to conducting certain types of studies. For example, reaction time-related studies can be difficult to operate as an online study.

iv. Other reality constraints include funds, technical know-how, poor infrastructure, etc. It is helpful to consider if the resources needed to participate in the project are available. Consider how they will contribute to the study, facilities, resources, capacity of collaborators, etc. For instance, they needed to train collaborators in statistical analysis in some situations.

v. There is a need to recruit collaborators for a large-scale study in Africa.

African psychology tends to have a different research focus compared to Europe. For instance, psychology is dominated by clinical research in Africa, with less focus on other areas of psychology.

Appendix II

1. Consultants take on studies to focus on

The central conclusion is that we should focus on a replication study instead of an original study due to the resources required to develop an original study. Also, we should consider a feasible and applicable research design in the local contexts. Lastly, these studies should be reproducible and ethically acceptable. See bullet-point summaries of individual consultants below.

- Study nomination and selection criteria should include nominating meaningful behavioral science studies, and as much as possible, we should avoid labeling effects as generalizable.
- When conducting the study nomination survey, the framing of questions for the survey should be direct/concrete in terms of study interest (e.g., ethnic conflict, poverty reduction), design (online or field experiment), and method. We should state which research area/topic we are looking to select effects/studies from and the research designs and methods.
- Simply put, the MLA study should focus on important issues that affect the public. Hence the study should consider effects or studies that can help resolve practical problems in these societies.
- To select a study, ensure that you consider studies that are relatable and of interest to the researchers and the general population. I suggest proposing the study and asking if they are interested, judging their commitment, explaining expected contributions, and sampling.
- As proposed, the study should be a replication study.

- Cross-cultural studies can be a workable type of studies to consider, such as a universal study of primary emotions (e.g., face expressions)
- To select a study, 1) consider that study is of interest to the local population, 2) select a study that addresses universality problems, 3) consider the research design that will be appropriate in the local context, and 4) what area of study works best in the local context. For Africa, clinical and cross-cultural areas of study could be the best bet.
- Finally, consider translating the measures and checking for the cultural equivalent of these measures.

Appendix III

1. Consultants take on a preferred study selection approach

Consultants opined that both approaches were workable. However, the crucial issue is the product of these approaches. They argue that the selected studies should represent the interest and expectations of local researchers regardless of any of the two approaches adopted. We decided to adopt the bottom-up approach because local researchers can contribute more to the selection process. A point-by-point summary of consultants' opinions is presented below.

- These approaches will help if the selection focuses on selecting studies of pressing issues for these countries. Investigators' priority should be to nominate studies/effects that focus on studies of pressing issues rather than approaches to selecting the studies.
- For a first step, survey local researchers on the "national pressing issues" the research study should focus on. Second, based on the pressing issues indicated, we would be able to narrow down to some research areas and interests. Third, local researchers can then nominate effects/studies/topics in these research areas.
- Both approaches are workable. However, the bottom-up approach would allow local researchers to nominate areas of research interest as a first step and then possible topics or studies of interest. For instance, the consultant described the Collaborative Replication and Education Project model as usable by adopting open science tools and practices for conducting collaborative replication studies. However, there is a need to encourage local users to adopt this approach.
- Moreover, most large-scale studies are resource-intensive, stretching the limited resources available to African researchers. For example, large-scale studies take longer to conduct than a regular study.

- Top-down will be more appropriate in situations where time resources are limited. However, adopting a bottom-up approach will allow local researchers to contribute at the early stages of the project.

Appendix IV

Table 20 Respondents’ study adaptability comments by continent for the extramarital affair study

RC	Comment
A	Adaptable
A	The materials could be adapted.
A	They are adaptable but not sufficient. The most important factors like family background, socioeconomic status/background, education and ethnicity could be very important predictors
A	Yes they can be adapted
A	The materials are adaptable in my country and the changed may include language and culture indicators in terms of marital relationship
A	The material should be adaptable I guess, because the topic always is that which raises interest among people and it is responsible whether or not marriages last!
A	The study can be adaptable to participants in my country but I can't give any comment on the materials since they are yet not available
A	The materials can be adapted and no changes are required
E	not sure
E	Probably adaptable, can't be sure without the materials.
E	I am sure these materials can be adapted. I'm only worried that we will not find the same effects in my country
E	It's difficult to assess the adaptability of the instruments, considering that we have no access to them right now. Given that they are based on "general views", I'm wondering to what extent these general views are easily adapted to the Italian context. For instance, what do the authors exactly mean by "home instability"? In sum, I'm not fully confident that these items could be easily and promptly translated/adapted
E	I think measures regarding extramarital affairs can be adapted easily. The reproductive health measures might take a bit more effort, as different health issues might be more or less common and relevant. I am not really sure about the home instability measure, there might not be that much variance, but I would need to take a closer look at the particular items
E	There are several measures that can be used without many cultural adaptations
E	Materials seem adaptable. I am not sure what is the content of variable home stability and it therefore is hard to judge its adaptability.
E	This is difficult to judge without seeing the materials, but I'm inclined to think they should be adaptable.
E	It seems that the measures would be valid only for some subgroups of populations. In general, we will have a low variance in the scores.

E	It probably could be adapted, but given the connotations this research brings it probably wouldn't be of interest for European public. Hinting that extramarital affairs are related with higher risk for reproductive health without taking in count education for example is something that could be seen as promoting traditional values and not objective finding
E	Adaptable
E	I am not sure if participants will say the truth about extramarital affair. maybe if become extramarital affair intentions
E	I think it can be adapted, generally speaking, marriage, especially early, is probably a bit less common
E	'Extramarital' affairs won't work; many couples are not married. In general, I'm not sure how this works with more diverse relationship structures.
E	I think the level of adaptation would be much too high to compare with the original
E	I said not interested, rated everything low for this reason, I cannot assess the study due to not knowing what the measures are, apologies
E	I would avoid stressing too much the correlation extramarital affairs - HIV
E	Difficult to say anything about the measures without having seen them. In general, it should be well possible to measure these constructs.
E	Yes, I think that the study materials could be adapted for participants in my country.
E	The materials can be adapted
E	not enough information to judge
E	Given that no items had been provided, it's not hard to judge. Thinking of health issues, we have to consider the country's base rates of STDs.
E	Will have to be very careful about data management and phrasing of questions around affairs but it seems possible
E	Seems suitable, but there will be social desirable responding.
E	I think it depends a lot on the instruments used to measure extramarital affairs and reproductive health.
NA	For the purposes of this replication project, I only have access to a university student sample. Because very few participants in this sample are married, it's not a suitable sample for examining this question.
NA	I am not sure whether the meaning of extramarital affairs is similar.
NA	I do not have enough information to answer this
NA	Possibly, but it's hard to tell. It depends on whether the questions measuring reproductive health were adapted to the most prevalent local reproductive health issues. The wording of the home instability measure could be an issue as well.
NA	It's not clear to me what the authors mean by reproductive health, and some questions on reproductive health might be very sensitive in the US right now given the recent ban on abortion in many states. I also wonder how honest people will be in reporting extramarital affairs, so how to ask about that would likely need to be pilot tested.
NA	It could probably work.

NA	If the materials are survey questions about extramarital affairs, reproductive health, and home instability, then they should be adaptable.
NA	I don't understand what they meant by reproductive health (operational definition), so it's hard to say. But I think that people in my country will not admit to an affair or reproductive health, or home instability (esp. depending on how things are worded)
NA	The materials seem likely to be adapted without too much difficulty.
NA	It is not clear to me whether the participant is asked whether they themselves engage in extramarital affairs, or whether they are reporting their perceptions of extramarital affairs. Both are interesting, but they each answer a different question. If it's the former, it would be hard to get a large enough sample of participants who states that they are currently in an extramarital affair, also it might not even be ethical to ask such a question...
NA	it is hard to tell without the actual measures, but unlike some of the other topics, this one is less familiar to me as a researcher, so my advice might be less useful anyway!
NA	It's difficult to make this assessment without seeing the measure but I from the general description I believe they should be adaptable.
NA	I would have to see the materials to say.
NA	Based on what I've read so far, I'm uneasy about the quality of the study and its materials.
NA	Probably, but not sure exactly what is included in the measure of reproductive health and if it might be somewhat co-linear with infidelity (e.g. if having multiple partners is an indicator of both poor reproductive health and also a component of infidelity)
NA	I'm not sure, but I would be more concerned with honest responding than material design. Although extramarital affairs in the US may be prevalent, it might be a problem to obtain honest responses from participants. Reproductive health would likely be okay to measure, but it could involve a variety of questions. Home instability might be difficult to capture in a similar way in the US due to cultural differences in familial structure and values.
NA	In general, it seems possible. I can't go back and look at the materials in this survey, so it's difficult to comment specifically.

Note. A – Africa, E – Europe, NA – North America, RC - respondents' continent

Table 21 Respondents' study adaptability comments by continent for the teenage pregnancy study

RC	Comment
A	Very adaptable
A	the materials are suitable, but response column should be three, i.e., agree neutral, and disagree, scored 3, 2, and in order of agreement.
A	Yes it is adaptable
A	The material is adaptable and does not require changes
A	The material is great and can be adaptable in my context. However, it is sometimes regarded as a taboo and a source of shame to the parents and guardian of the affected girl. To me I will prefer "Feel Shameful" for "Devastated" in item 3 of instrument 1.

A	The materials can be very much adapted to participants in my country
A	The materials can be adapted quite alright, but not worth further replication in my country, as it does not constitute a significant social problem worth exploring further
A	the instruments may be adapted
E	harder to adapt in my country, due to cultural and social differences. for example, "send to mission home for care" is not an option people would really understand here. also, "stop her schooling" would not really be an option.
E	Major changes would be needed, the items seem to be culturally specific.
E	Materials can easily be adapted, but teenage pregnancy is not a general problem in Denmark
E	I'm not fully sure that these instruments could be adapted to the Italian context, without introducing some elements which are not present in the original instruments. For instance, "mission home" are not present in Italy. Moreover, I'm not sure how I could adapt the reference to the Western education.
E	Adaptable but "Western education" would need to be changed
E	I can imagine that people might feel a bit offended, ethically, by these questions.,especially given that the response scale is binary (rather than giving more options)
E	These materials can be adapted, but I think we should revise the response scale for a continuous one (e.g., from 1 to 7). Otherwise, responses may lack some variability and the analyses will be limited
E	Although materials are largely adaptable to my local context, some items from the scale Attitudes of parents towards teenage pregnant girl might not be applicable in my context.
E	I think these materials are adaptable. Not all response options may be equally common, but they can certainly occur.
E	Yes, the most of the items are suitable, but some are not (e.g., "stop her schooling" is not relevant for our country).
E	I think that this material could be adapted to some extent. However, attitudes and actions related to them (especially practices regarding showing those attitudes are somewhat specific in the original version. I believe that somewhat similar list could be achieved but it probably wouldn't be equivalent)
E	Adaptable
E	I love the study!
E	Recently, there is a lot of discussion around abortion. I think the material is adaptable
E	While maternal poverty is a debatable subject in my country as well, the implications are by far less severe, compared to the situation, described in the abstract
E	Very little teenage pregnancy here.
E	It is adaptable - - with some changes. Some items are a tad too extreme (Stop her schooling/is a deviant)
E	They can be adapted, however the wording would need to be changed for some of the items.
E	adaptable. I would erase the last sentence in Table 1 because not pertinent in a Western Country like Italy, where teenage pregnancy is rather the effect of a lack in education
E	Dichotomous response formats are almost always poor and should be adapted to ordinal or interval level. "... result of western education" will have a different meaning in Africa than in western countries. "stop her schooling" is not allowed in many countries and therefore not an option. Some options are just not realistic in my country and some other

	rs are vague. Not sure comparing these situations across countries provides many insights.
E	Some of the attitude-items are not really applicable to my country: Sending teenage pregnant girls to a mission home or stopping her schooling would be extreme responses. These items would need to be replaced by reactions that are more common in my country (e.g., sending her to a shelter for teenage mothers)
E	Materials are adaptable
E	"give her out in early marriage" and "teenage pregnancy is the result of western education" would have to be adapted.
E	I think that the questions in Image 1 are pretty easy to implement. Especially, "Teenage pregnancy is the result of western education" could be phrased in way that we could use it for cultural comparison. The questions in image 2 are tricky, because it's not possible to stop schooling the girl or to send her to an early marriage.
E	Easily adaptable. Some changes would be necessary, such as the item "Give her out in early marriage" (could be changed to force her to marry, although it's not allowed to marry under 16 years in Portugal), and the item "Send to mission home for care" ("Send to shelter for teenage pregnant women").
E	This might be considered a little 'strong' so would need rewording, particularly those on the second image where most items would need re-writing or reconsideration (e.g. send to mission home for care)
NA	The item about giving her out in early marriage could be adapted to "encourage an early marriage"
NA	There are rarely teenager pregnancy in total and many of the options would not apply at all.
NA	Binary scale is odd here, but usable. Some language tweaks might help. Instrument 2 has no instructions
NA	I don't think these would be adaptable. There are stigmas towards pregnant teenagers, but it's generally not acceptable to be outright about it and call them deviant in many parts of the country, and many of the other statements (not having a future, being devastated, result of Western education) would be difficult to apply. Additionally, parents have less control; they couldn't stop schooling (that's illegal), send her to a mission (boarding school, but those are uncommon), give her out to marry (they could pressure her to marry, but that's different). This would show very strong cultural differences, which is valuable, but we'd likely get floor effects here. A more implicit measure would be needed to capture views on pregnant teens.
NA	These could be easily adapted. The only one that isn't relevant for the US is the item on western education.
NA	This can be. I'm not sure what a mission home is, so that would need some clarification. Referring to western education wouldn't be so relevant either; maybe "liberal education"?
NA	These are adaptable but some items are culturally specific and would not apply in the US, or are awkwardly worded for American participants: "send to mission home for care" "give her out in early marriage"
NA	yes, several wording changes would be needed. I'd also suggest more detailed questions, since "support girl to have baby" for example, could mean a lot of different things. Western education would also need to be changed
NA	I think the materials could be adapted, but some items are less culturally relevant or would (likely) be infrequently endorsed. For example "result of western education" "give h

	er out in early marriage" "send to mission home for care" seem less relevant. Additional items suited to the cultural context may need to be developed, but that would limit the potential for direct cross-cultural comparisons.
NA	I think that a lot more measure would be needed, given that this study is cross-sectional. Also the nature of the scales (Agree or Disagree) would have to be presented in a continuous fashion rather than dichotomously.
NA	the term "mission home" is unfamiliar to me, and may be to other US participants. Overall, the methods seem reasonable, though I would generally prefer a scale rather than dichotomous option, but perhaps in an African context, 5 point scales are not common, so I understand why a dichotomous response might be appropriate.
NA	I think the materials can be adapted. The "Teenage pregnancy is the result of western education" item may need to be changed, or perhaps it would be interesting to keep it in and see how westerners respond and compare it to this study.
NA	These materials would be difficult to adapt to the United States because the education is western and the culture is relatively loose when it comes to norms about sexuality. Many parents who would be appalled that their daughter got pregnant outside of marriage would rather the daughter have the baby and continue to accept her in their family, rather than encourage her to have an abortion. Stopping schooling while she is in high school would be illegal here, and parents do not give away daughters in marriage. With that said, in Image 1, all but the last question could be usable. In Image 2, the last two items could be usable.
NA	Yes, I think they're adaptable but it would be challenging. E.g. it seems like "deviant" is used in these materials in an idiosyncratic way that feels outdated for the US. "Western education" would mean something different to a US sample. Sending people to "mission homes" is not really a thing in the US.
NA	It seems like each of these instruments is meant to cover the spectrum of most common/applicable perceptions and attitudes for the parents. While it would be possible to adapt these items, the question is if these items cover the most common perceptions/attitudes about teen pregnancy in other countries/cultures. That seems unlikely, and presents a challenge.
NA	I think the attitudes of parents toward teenage pregnant girl would need to be substantially reworked for a US context, but I don't have ideas of how that might happen. I don't know that I would call them attitudes though. The last perceptions item would have different implications across cultures, probably.

Note. A – Africa, E – Europe, NA – North America, RC - respondents' continent

Table 22 Respondents' study adaptability comments by continent for the civic engagement study

RC	Comment
A	They are ideal for Nigerian context
A	Control as Features of a table is not clear. I suggest this could be changed to agreed profiles of success measures for all to read. The control I, should be chosen from perceived achievers on these measures. Scores of other groups could then be measured as comparable performances to know differences
A	It can be adapted but I fear that a lot of data may not be useful in the end
A	The materials are adaptable but I think inclusive participation and organizational citizenship should be included in the study

A	They're adaptable
A	The materials can be adapted in my country because it portrays the same cultural meaning. My opinion is that the social norm intervention may be modelled according to the country of the collaborator and not using Uganda for all the participating countries. This I think will have a more direct effect or home feeling to participants when it concerns their individual countries. Example - "participants were narrated a story of an ordinary citizen who successfully achieved change for the better in her community in "Nigeria".
A	The instruments have to be adapted
E	Adaptable
E	It will be very difficult to adapt. Civic engagement is extremely diverse in Germany and there might not be one straightforward dependent variable that fits most participants.
E	This materials are adaptable to my country, Denmark. However, engagement in civic issues might be very different in Denmark due to 1) extremely high trust in government and 2) the construction of the Danish welfare state.
E	The materials should be somewhat adapted. For instance, it should be made clear what we mean by "community" and "leader". It seems that these two key feature should be very different between Uganda and Italy. Moreover, it is not clear to me what constitutes the outcome variable (i.e., intention to participate?)
E	Adaptable "community in Uganda" needs to be changed
E	I think these materials are easy to adapt
E	I think the problematic with these materials is their abstractness. Unless we get specific examples in the community, it's very hard to understand these manipulations, from a participant's perspective
E	Materials seem adaptable.
E	These materials seem easy to adapt. One possible concern to consider is that the social norm intervention also contains an element of efficacy as the change attempt of the citizen was successful. In that sense the manipulation may not be entirely clean.
E	The materials look short, but the main ideas are suitable to use.
E	I think that materials should be only translated
E	Easy adaptable. Translation only
E	Adaptable
E	It applicable, the major question is whether the manipulation is strong enough to influence behavior. I doubt so, unfortunately
E	I think the material is adaptable. I believe a corruption intervention will be also relevant
E	Im not sure if the "business as usual" style is commonly used in my country of research
E	I think the outcome measure will be tricky to standardise across societies; community meetings are not that big of a thing here.
E	This is out of my expertise area but I think this could be adapted. Probably the idea of "leader of community" is the most difficult.
E	Yes
E	I think the materials are suitable as they are. Of course the story narrated should be contextualised (not referring to Uganda) in order to be acceptable as an example in a Western country
E	Unclear what "business-as-usual" rhetoric is in different countries, self-efficacy intervention was apparently not tailored at the specific behavior in question, u

	nclear what a "community leader" is and whether most people ever talked to one, story of *one* other citizen ist *not* a descriptive norm.
E	Yes, I think materials are generally adaptable, but the understanding/relevance of "Community" might be different in my country.
E	The materials are adaptable
E	Not enough information to judge how adaptable the materials are
E	The materials should be pretty easy to adapt.
E	It wasn't clear to me what kind of information was provided in the placebo condition; thus, it's difficult to make suggestions. Self-efficacy and social norm intervention conditions are easily adaptable. My biggest concern is the dependent variable: attend a meeting and how to operationalize it. It was also not clear how the sample was selected: convenience sample? Students? Despite the importance of this type of studies (measure actual behavior), usually, it's extremely difficult to recruit participants when the study requires a certain level of effort (e.g., going to the lab; attend a meeting). With students it might be possible to operationalize this type of study.
E	The content would need changing but the scope of the intervention types would be possible.
NA	These seem adaptable by changing the reference to Uganda to the United States or, ideally, a more local community.
NA	Replicating a "null-result" is problematic.
NA	Norm changed from Uganda to local country
NA	I think they can be adapted. Most participants in my subject pool won't have access to leaders, so that part of the self-efficacy intervention prompt may not apply. Without seeing the "business-as-usual" rhetoric prompt directly, it's hard to tell how well that would fit but it would likely need to change to reflect "business-as-usual" in local cultures. Perhaps a local version and one based on the original prompt (creating five conditions) would be useful. For the social norm intervention, it's likely that the culture and character would need to change to make it more local, but again a comparison here between a locally-framed story and the original could be interesting though creating too many conditions could be a problem.
NA	The control and self-efficacy interventions could work as is in the US, but the placebo and social norm intervention would need to be adapted. I think the social norm intervention would be easy to adapt, but I'm not quite sure what the placebo intervention would include.
NA	The country would need to be changed, but otherwise I think it's hard to judge from these materials. Sorry. It's also unclear what the DV is.
NA	These might be adaptable to the general population but may not seem super relevant to college students who do not have experience with civil society organizations.
NA	this study would really need to be done in the lab, so that would make getting participants harder. But otherwise, it seems adaptable
NA	Adapting the materials should be relatively straightforward, though what constitutes "business-as-usual" would need to be decided.
NA	I found the conditions to not be as precise as I expected them to be. Also the control conditions appeared a bit too random for what I am used to seeing as control conditions. Given the fact that the study is also not published, I have less confidence in this study.

NA	I think that in general, these materials and ideas can be adapted to the US context. One concern is that with lower US civic engagement in general, it may be difficult for people to complete the key condition (self-efficacy), which asks for examples when participants have said something to a leader in their community (which might not be a common experience).
NA	I think these materials are easily adaptable.
NA	A lot would depend on the specific wording of the instructions, so it's hard to say from Image 1.
NA	These materials seem like they could be adapted without making many changes.
NA	Sure, these could be adapted. But if none of the interventions was effective... what is the goal of the study? To see if they are similarly ineffective elsewhere? Or is there some reason to believe they would be effective in another cultural context?

A – Africa, E – Europe, NA – North America, RC - respondents’ continent

Table 23 Respondents’ study adaptability comments by continent for the self-disclosure study

RC	Comment
A	Very relevant to Nigeria
A	It is adaptable
A	This material only covers student population only. It will be better if it covers literates both learned and educated people. It should cover places of residents and Status.
A	Yes it is adaptable
A	The materials are adaptable and slight changes especially in length of words should be more concise
A	It is adaptable and to me I'm ok with the material
A	The materials can be adaptable to participants in my country
A	The material can be adapted to be applied to users of twitter
A	The instruments are suitable to our context. We need to translate it only
E	Very adaptable
E	The materials should be adaptable. Aside from translation, no changes are necessary.
E	These materials can surely be adapted to my country, Denmark. However, I fear that Generation Y in my country are no longer using Facebook to such a high degree anymore.
E	I don't anticipate any problem in adapting these items to the Italian context.
E	Adaptable without any changes
E	I think young people in my country don't use Facebook anymore, so it might be better to ask about e.g. Instagram which is more common. Besides this issue, it should be adaptable
E	I think the materials can be used in the study as is
E	Materials are complete, adaptable to local context of my country.
E	I think these items can be used pretty much as is. No changes seem necessary.
E	Yes, but it needs another social network, because Facebook is not popular in our country.
E	These materials can be adapted for Russian context through changing Facebook to Vkontakte site (the most popular network in Russia)

E	I think that this material is easy adaptable. Probably, given that Instagram and Tick-tock are most popular social networks and Facebook is seen as something for "older people", referencing to Facebook should be replaced with referencing to Instagram
E	can be, but those questions are weird, it is a normal thing to use FB and maybe those questions would be strange here in Europe
E	One can also add sociodemographic indicators (social class), and other important indicators, like self-esteem.
E	I think that the materials are adaptable and fits to cypriot population
E	It definitely is adaptable, the problem is, that current students barely use Facebook in this manner anymore. I would only favor this study, if other networks are used, such as Instagram and TikTok
E	Facebook has been rapidly using popularity among younger generations. I do not believe a focus on Facebook is appropriate.
E	This is the only study (so far) I don't think it is applicable. For one "generations" do not exist. See work by Sociologist Cohen, Philip.
E	yes they can be. Facebook is not as "big" in my country anymore though, many people use Instagram or TikTok but if we highlight Facebook users it should work.
E	The materials are adaptable, but in my country young people aged 18-24 do not use Facebook, rather they focus on Instagram or Tik Tok. Therefore I would suggest to adapt queries to these other platforms
E	Unclear, what "intimate" means. / Importantly, in my country, few young adults use Facebook. The majority does not use it. They use Instagram.
E	Facebook is not be the most-used / influential social media platform for the age group of people between 18 and 24 in my country. Should be change to Twitter, which is mostly used by people at that age.
E	The materials are adaptable
E	possible to use as is
E	The items can easily be adapted.
E	All the measurements are easily adaptable (only translation needed).
NA	If the question is about adults in the U.S. aged 18 to 24, I don't think a high proportion in this age range is on Facebook. If the question is about adults in the U.S. born between 1986 and 2005, I still think a high proportion is on Facebook.
NA	Students in Germany do not use Facebook anymore
NA	Facebook wouldn't be the right social media tool anymore
NA	From what is shown here, this looks like it would be adaptable. The main issue is that participants in my subject pool don't use Facebook that much. If we had funds to collect people online, then we might get more people who use Facebook. Participants in my subject pool use Snapchat, Instagram, and other social media platforms. For this to work, we'd need to change the prompts to be about social media platforms in general, ask participants about multiple platforms, or have participants choose from a list of platforms and then complete the questions about the one they use most. For this last option, we could look for differences across platforms and potentially collapse across platforms to look for general patterns.
NA	These materials could basically be used as is
NA	I'm not sure Facebook is the key social media site anymore?
NA	These are adaptable and don't need modification

NA	yes, definitely adaptable. I would consider changing the word "intimate" as I think college students currently have a different meaning than what is meant by the scale
NA	The materials seem to be usable without modification.
NA	Pretty straightforward measures, and relatively easy to adapt.
NA	My biggest concern is the datedness of Facebook as the modal social media site, and the potential variability across cultures in the extent to which any given social media site is "the" social media site for the participants studied.
NA	I think the materials are adaptable and do not need changes.
NA	These seem easy to adapt to the United States - really, these questions could be used as they are.
NA	Yes, these materials could probably be used as-is.
NA	Sure, but in some countries facebook use among college students is very low-- they've moved on to other platforms. I'm not certain that the platform-specificity fo this project is meaningful any longer.

Note. A – Africa, E – Europe, NA – North America, RC - respondents’ continent

Table 24 Respondents’ study adaptability comments by continent for the condom use negotiation study

RC	Comment
A	This materials can be adopted in Nigeria.
A	it should be adaptable if culturally validated for use in African society. Material 5 should address the number of times rather than the frequency . The continuum if adapted should go from n one all, to everyday. Try to build in negative response items to detect yes or nay sayers, i.e. lie detector questions .
A	Easily adaptable
A	It can be adapted in my country
A	To me the instrument/ material are adaptable in my context.
A	The materials can be adaptable to some participants in my country and may not be for so me others. I think some factors may influence this type of study in my country, example religion: religious people may not want to take part in this type of study. Secondly, the i mage 5 may not be culturally adapte because it is too direct and may jeopardize the integrity and morality of the researcher. Students may think the researcher supports their sexual activities and it may encourage them to get more involved in such act. Infact, the study may be modified to be based on perception.
A	No changes required
A	The instruments are suitable
E	Seem adaptable with minor changes, if any
E	The materials are adaptable.
E	Materials are easily adaptable to my country, Denmark. However, birth control pills like p- pills or hormone spirals are extremely accessible in Denmark and condoms are very cheap and accessible meaning that the problem is not very prevalent in Denmark

E	The materials shown could be easily adapted to the Italian context with little changes, especially if they will be used with young adults (i.e., not minors). I don't anticipate specific problems in translating the items to Italian. No specific adaptation will be required.
E	Adaptable without any changes
E	Easily adaptable
E	Measure can be used as is
E	Materials are fully adaptable to my local context.
E	These measures are easily adaptable. Few changes (if any) would be needed.
E	Yes, it is completely suitable.
E	Adaptable. Only some sentence constructions related to the nature of the language should be made
E	a little weird as in my country it is very obvious to use condoms
E	there is no sexual education at schools . Thus, I believe it would be interesting to hear their opinions
E	I think I would suggest to reduce the heteronormative assumptions of the material and refer more to "partners" and frame "the pill" as "even if alternative measures of birth control are applied"
E	I don't think the norms measures are very good (mix descriptive and injunctive norms).
E	Yes, it can be adapted.
E	All good
E	These materials are suitable as they are
E	Yes, adaptable for participants in my country.
E	Yes, I think the materials could be easily adapted for participants in my country.
E	Materials are adaptable
E	I would adapt most of the items to make them gender-neutral. Instead of "girlfriend/boyfriend", I would use "partner".
E	That should be rather easy.
E	All these materials can be easily adapted for my country (Portugal).
E	Some changes might be necessary e.g, the reference to 'my age' in the first set of questions is awkward, but otherwise this looks adaptable
E	"person my age" and "boyfriend/girlfriend" needs to be adapted when the study is conducted with adults; in Germany it would be necessary to differentiate between condom-use in long-term and short-term relationships
NA	These materials seem adaptable. Maybe change 'boyfriend/girlfriend' to 'partner.'
NA	I am not sure but it seems you consider Gender as two-dimensional (female, male) and miss other genders. Further, which type of sex is meant? Only male-female? Or all types of sexual contacts? Why do you refer to sex with boy/girlfriend only?
NA	I question the unidimensionality of some of these instruments
NA	I think they can be adapted. I'd recommend taking out references to age (a person of my age) in the Attitudes Towards Condom Use questionnaire. When adapting this for an older population, the Intentions to Condom Use and Self-Reported Condom Use questionnaires probably won't work for people in long-term stable relationships. Condoms are more likely to be expected for new relationships or casual sex. Rather than focusing on a specific time frame, it might be better to use hy

	pothetical scenarios where someone would be having sex with a new partner. The participant could answer questions about whether they've used condoms for a new relationship or casual sex in the past, and if they would in the future.
NA	These measures could essentially be used as is
NA	These would work. It might be useful to figure out how to best use them for non-heterosexual people (some items seem to be primarily aimed at heterosexual people)
NA	These are adaptable; no changes are required as far as I can tell
NA	yes-- take out british spellings, references to boy/girlfriends (to include gender neutral option), don't call women "girls"
NA	The materials seem to be suitable with minimal modifications.
NA	I think they can be adapted, they are pretty straightforward. The wording might need to be simplified a bit, but other than that they are fine.
NA	Methodologically, this seems fine, though TPB researchers often argue that the assessment of these variables need to be heavily pretested in a population-specific way. For example, perhaps for some groups peer norms are most important, but for other groups parental norms would be. Also, the analyses in the main paper seem to examine mean differences only, not differences in effects of variables on BI as a function of gender. I'd be very curious to examine differences in these slopes as a function of culture.
NA	These materials can be adapted. Wording will need to be updated to reflect that the replication will be with adults rather than children (e.g., taking out "girl" and "boy" wording). Wording will also need to be updated to be less heteronormative (or alternatively we would need to restrict the samples to only those who identify as heterosexual).
NA	These would be easy to adapt for use in the United States.
NA	Yes, these could be adapted. Might consider changing boyfriend/girlfriend to romantic partner or sexual partner or something like that.
NA	Yes, could be adapted. The only issue I see is that these items apply to heterosexuals, so there would need to be some branching for those with different preferences
NA	I think the materials are fine, but I also think that replicating this study of high school students with adults would not really be a replication

Note. A – Africa, E – Europe, NA – North America, RC - respondents' continent

Appendix V

Table 25 Respondent’s additional comments on the extramarital affair study by continent

RC	Comment
A	This type of study is urgently recommended for Africa continent for obvious reasons.
A	This study might be sensitive to cultural relativism in my country as marital relationship is something Africans hold sacred due to cultural affiliation
A	It's a good research topic
A	the instruments will be adapted if needed
A	It could have been nice to get the measures used.
A	This being a survey study it will more replicable online.
E	I anticipate that an Ethical Committee could be anxious about collecting information on extramarital affairs and STDs. Practically, I presume that individuals would be more willing to disclose personal information online, rather than in vivo (i.e., paper-and-pencil format)
E	The online sampling of couples will be a bit more effortful, than for individuals, especially for such topics
E	The idea that extramarital activity by and large correlates negatively with relationship factors such as stability strikes me somewhat trivial. / I understood that not individual participants, but couples participated in this study, although this was not reflected in the data analysis. Recruitment of couples is considerably more difficult than recruitment of independent individuals.
E	I don't foresee a problem in getting ethics approval for this study, but there is an ethical problem in the self-selection of the sample. As the original article says, extramarital affairs are taboo in many countries, and people who have had affairs may decline to answer a questionnaire about them even if data collection is anonymous.
E	Interesting study!
E	See in previous answer. What was the direction of the correlation?
E	Concepts don't seem clear.
NA	I'd be worried that people would see this study as shaming people who do not have traditional marriages.
NA	It's hard to evaluate this one without knowing what the measures look like
NA	The paper didn't have a lot of detail and I'm a little confused at the extramarital affairs and positive significant correlation with reproductive health leading to conclusion that affairs are related to STDs (perhaps the label was supposed to be reproductive health issues or something similar?). Since we don't have the measures it's also not clear how extramarital affair is defined and if this includes consensual non-monogamy.
NA	It seems that it would be difficult to get participants for this study if they were to be reporting on their own behavior.
NA	The study seems poorly designed and described and I'm not sure a replication is a good idea.
NA	I'm not sure that extramarital affairs are highly prominent when people think about their concerns in the US.

NA	It may be somewhat more difficult to recruit people only in committed relationships.
NA	I think the study could be interesting to run. The materials would need to be reviewed and the reason WHY would need to be explained a bit better as well.

Note. A – Africa, E – Europe, NA – North America, RC - respondents’ continent

Table 26 Respondent's additional comments on the teenage pregnancy study by continent

RC	Comment
A	Great work
A	It is germane in a high fast growing population country like Nigeria
A	This study can help address parent to child bonding and communication that will mitigate some of these issues!
A	If the target population remains the rural population, they may not be able to participate in an online survey due to technological challenges.
A	Most of the girls that are victims come from poor communities and therefore, may not have smart phones, whilst the questionnaires have to be administered to them
A	If the participant must be from the rural community, it would be difficult to conduct the study online. If it concerns the general population, it is feasible.
E	I would suggest changing the response scales so that for each option/item participants answered on a Likert scale (e.g., 1=strongly disagree, 7=strongly agree).
E	I'm not sure which psychological effect is being demonstrated here. The abstract sounds more like a survey decoded by demographic data.
E	if the sample is adults 18+
E	Questionable relevance in my country (teenage pregnancy rate is at about 1.3 %; also - not taboo to the extent that is suggested by the instruments). It is feasible and ethically acceptable though and could be OK if adapted
E	Do you expect to conduct this study just with people who are parents?
E	See above. In addition to my previous response: the items could be used in German, but could be a bit harsh (“deviant”)
E	Accessing rural populations online would be a challenge.
NA	I think the study would require radical changes to the materials to become acceptable.
NA	Would parents specifically need to be targeted, rather than a general population of adults? It would be possible to have a sample of parents, but would require some kind of screening.
NA	Identifying an appropriate "core" set of items that are relevant across cultures seems to be the main challenge for this study.
NA	Basically, simply replicating this study would not be scientifically novel, as there are some measurement concerns, and it is not clear what the empirical contribution is. More measures are needed to justify replicating this study.
NA	It would be difficult to find parents if the contributing institutions work with student samples. This could be a hard topic for which to use snowball sampling.
NA	I hope attempted replications of this study would include more than binary response options for the DVs of interest.
NA	I'm not completely sure what 'the effect' would be to try to replicate for this study? Same level of endorsement for the same specific attitudes about teen pregnancy? Why?
NA	I think prevalence of the abortion issue in the US might complicate this study.

NA	It may be somewhat more difficult to recruit parents specifically for a study such as this .
NA	I believe the study can be ethically acceptable if the background includes why there is a centering on girls, specifically and what this study aims to bridge gaps on within the population.

Note. A – Africa, E – Europe, NA – North America, RC - respondents’ continent

Table 27 Respondent's additional comments on the civic engagement study by continent

RC	Comment
A	Commendable research
A	There should be an initial brief, adequate and generally informative to selected participants in order to comprehend the experimental nature of their individual roles. Give time interval between this phase and research proper.
A	online studies may be limiting because the participants should own smartphone or have access to a computer
E	This study seems to be better suited for a field study design. It is possible to do it with follow-up online surveys, but I also think, the dependent variable will be an issue.
E	It feels like a substantial cultural adaption would be necessary
E	The DV (one-shot measurement of one particular behavior at one particular time) is necessarily unreliable and of limited usefulness to reflect an effect of the intervention. Also, one-shot, brief interventions are unlikely to change anything on a broader scale. Not sure of the value of this study.
E	I'm not sure about the authors' conclusion. Basically, if there is no significant effect they shouldn't conclude that the interventions had the same effect. You can't do this with NHST.
E	Ethics might be concerned by the scope of the interventions so these would probably need to be developed prior to securing ethical approval.
E	I imagine that the outcome measures and social influence manipulations may be very culture-specific. A failed replication may say more about those details than about the basic psychological effect.
NA	I'm guessing the measures were a survey about intentions to act or something like that, in which case this would be feasible so long as they aren't about addressing issues around sensitive topics.
NA	I didn't really have enough information to judge feasibility.
NA	The dependent measure, as briefly described in the report, seems less adaptable for an online context, but overall, the other elements seem easily doable in an online context. If this was adapted in a way that was doable online, I'd be very interested in this project, given my interests in persuasion and social influence more generally.
NA	It's hard to say much about this study - I'm not seeing a lot about why it would be interesting or specifics about the method.
NA	Getting the behavioral follow-up measure could be challenging. Did they measure any other variables besides attendance at the next community meeting?

NA	The dependent variable isn't described very well. I'm a bit unclear about the goal of replicating this study-- is it to confirm that these interventions are not useful? Does that contradict an important theory or prediction? If so, cool... but if not, I'm not sure if I understand the reason to replicate social interventions that didn't influence behavior substantively.
NA	I am highly skeptical that norm interventions affect behavior, especially in large groups. The smaller the groups I think the more likely that there would be effects of norms
NA	Yes, see my last comment, with concerns about manipulation checks.

Note. A – Africa, E – Europe, NA – North America, RC - respondents' continent

Table 28 Respondent's additional comments on the self-disclosure study by continent

RC	Comment
A	Good work. Keep it up
A	Language usage is extremely important for respondents reactions. Please exterminate foul languages or words.
A	This study is apt!
A	Easy to do online study but not so relevant to policy
A	It is a contemporary study suitable for my country.
E	I think the study may not be that interesting because a lot of users leave facebook for other social media.
E	This study is a bit less exotic than some of the other ones, but it is great from the point of view of cross-cultural comparability.
E	See my previous note. I wouldn't advise this study to be part of the ML
E	I think the theoretical reasoning of this study is rather poor.
E	Ethics might have concerns if we have to access other people's FB pages etc. but looks reasonable to me.
E	As far as I know, Facebook is becoming less popular in my country at the moment. So there may be readers who question the everyday relevance of this study. Perhaps a variety of social media could be investigated.
E	I have explained the position above.
E	Perhaps as extension: self-disclosure to others outside Facebook
NA	just about the age/generational difference in Facebook use within the United State
NA	I'm not sure it would make sense to look at this question among a sample of university students, because I think they tend to use other social media sites more than facebook. However, it could be examined among an adult sample.
NA	The social media landscape has changed significantly in recent years. Depending on the ages of the intended sample, Facebook may be a relatively unpopular social media platform today compared to the original study.
NA	Of the four studies I've assessed so far, this is by far the most feasible.
NA	I fear that this study would be challenging for us to conduct, mainly because it is challenging to reach a larger sample of minors.

Note. A – Africa, E – Europe, NA – North America, RC - respondents' continent

Table 29 Respondent's additional comments on the condom use negotiation study by continent

RC	Comment
A	A good research
E	German regulations about personal or sensitive information especially with underaged students can be strict. Recruitment of this sample could also be more complicated, although it certainly is doable.
A	It will be good for policy making on sexual health behaviour and population control
E	It might be hard to reach 600 participants or more. Gender imbalance would probably occur in sample.
NA	This could easily be done in an online sample of adults in the US, or among a university sample
A	I think the setting whether rural and urban adults would affect the results and it should be considered
NA	I would think this has been done in the US already b/c there are so many TRA/TPB studies. Might be interesting to see if it's changed now that abortion isn't easily accessible for most Americans
E	This is a fairly well-studied topic; it's not clear to me why a multi-country replication is necessary.
NA	The authors of the original study may have conflated gender with biological sex. In their Instruments section there is no indication of how gender was measured/determined, and while the term gender is used throughout the paper the participants' data just show Male and Female numbers. A replication would need to have more clarity on how gender/sex is being defined and determined.
NA	I wonder how informative replications will be if they focus on adults? It seems to me that part of what's interesting and important about this study is that it describes adolescents.
A	I'm ok with the topic and what it intends to measure.
A	it is a sensitive theme regarding the Rwandan cultural
NA	As before, I don't think that replication of this study of high school students with adults would be a replication of the same effects
A	Although the target audience is adults in your replication, the most at risk group remains the adolescent group who may be in high school at this stage.
A	This being a study targeting adolescents the ethical issues are complex in my country.
E	My main concern with this study is its focus on gender differences, which would require a sample that is representative in terms of genders, neglects some subtler points about gender identity and sexuality, and may ultimately not be of much interest. What would we really learn from a successful or failed replication of the effect that men feel more in control of condom use than women, for example?
NA	Unfortunately, we don't have easy access to high school students in our country. And, to be honest this would be the only acceptable sample that I see unless the scope changes to college use.
A	Some people can be resentful to answer questions online
E	The study is feasible/acceptable, but has questionable relevance

E	We would have to clarify that only sexually active individuals take part, and probably only heterosexual individuals?
---	---

Note. A – Africa, E – Europe, NA – North America, RC - respondents' continent

Appendix VI

1. Teenage pregnancy study results for full sample

1.1 Agreement proportions for the full sample.

Without a specific hypothesis to test, we estimated the agreement proportions across all regions for individual 13 items and their CIs and tested whether the CIs overlap between the original and the replication samples. We present our findings below in Tables 30 and 31, the estimated probability for the 13 items estimates on attitudes towards teenage pregnancy, and their predicted probability CIs for the full sample, and the original calculated CIs used for replication test if those overlap with replication predicted probability CIs.

We excluded the North America sample for the subgroup analysis because the sample of parents aged 35 and above is too small and cannot be comparable to the African and European subgroup for this demographic. However, in this auxiliary analysis, we included the North America sample and tested these focal effects using the full models (see Table 9). We observed that 11 out of the estimated probability for the 13 items were significant for the entire participants (see Table 30). The participants' attitudes towards teenage girls were positive. Participants disagreed with negative notions and actions (e.g., the pregnant teenage girl does not have a future or her parents have failed), and agreed with positive ones (such as supporting the girl). For Item 3 "Teenage pregnancy makes parents devastated and Item 13 "Parents should force the man/boy responsible for the pregnancy to accept responsibility of her care" that were not significant, participants were indifferent about their agreement on these attitudinal statements. These focal effects and replication tests were consistent with the subgroup findings (see Tables 14 and 15). However, we observed that the by-site variability between and within regions differs in respect to the variables for the full sample.

Table 30 Teenage pregnancy replication results for the full sample

Item/Statement of test	Estimated probability for $b_{intercept}$	95%CI predicted probability for $(b_{intercept})$ Lower, Upper	Original study 95%CI for agreement Lower, Upper	Replication decision	$sd(b_{intercept})$
<i>Item 1: A pregnant teenage girl is a deviant</i>	0.189***	0.141, 0.248	0.904, 0.960	Did not overlap, Not replicated	1.09
<i>Item 2: A pregnant teenage girl do not have future</i>	0.048***	0.039, 0.059	0.797, 0.879	Did not overlap, Not replicated	0.47
<i>Item 3: Teenage pregnancy makes parents devastated</i>	0.542	0.483, 0.599	0.819, 0.897	Did not overlap, Not replicated	0.72
<i>Item 4: Teenage pregnancy means that parents have failed</i>	0.096***	0.081, 0.112	0.815, 0.894	Did not overlap, Not replicated	0.45
<i>Item 5: Teenage pregnancy results in stigmatization</i>	0.771***	0.734, 0.805	0.859, 0.928	Did not overlap, Not replicated	0.59
<i>Item 6: Teenage pregnancy is the result of western education</i>	0.144***	0.127, 0.164	0.651, 0.753	Did not overlap, Not replicated	0.39
<i>Item 7: Parents should drive girl away from the family in anger</i>	0.016***	0.012, 0.023	0.275, 0.379	Did not overlap, Not replicated	0.68
<i>Item 8: Parents should stop her schooling</i>	0.031***	0.024, 0.039	0.442, 0.554	Did not overlap, Not replicated	0.56
<i>Item 9: Parents should give her out in early marriage</i>	0.044***	0.032, 0.062	0.574, 0.682	Did not overlap, Not replicated	0.89
<i>Item 10: Parents should abort pregnancy</i>	0.149***	0.118, 0.187	0.118, 0.199	Overlapped, Replicated	0.80

Increasing the generalizability of psychology by including Africans

<i>Item 11: Parents should send her to a maternity group home for pregnant teenage girls for care</i>	0.209***	0.173, 0.249	0.644, 0.747	Did not overlap, Not replicated	0.79
<i>Item 12: Parents should support girl to have the baby</i>	0.835***	0.795, 0.868	0.521, 0.631	Did not overlap, Not replicated	0.79
<i>Item 13: Parents should force the man/boy responsible for the pregnancy to accept responsibility of her care</i>	0.482	0.438, 0.526	0.801, 0.882	Did not overlap, Not replicated	0.52

Note. A – Africa, EU– Europe, NA – North America, * $p < .05$, ** $p < .01$, *** $p < .001$

A summary table showing the estimated probability of the focal effect for the ($b_{intercept}$) for overall sample, and their predicted probability CIs and standard deviations. For the replication decision, when we compared the replication CIs predicted probability and original CIs. Only Item 10 CIs for original and replication overlapped, hence were replicated.

1.2 Region comparisons for agreement proportions for full sample

To compare the regions' proportion agreements for these 13 items, we fitted dummy-coded indicator variables for the comparisons between Africa and Europe as well as between Africa and North America as fixed slopes in our model (see Table 9) using Africa as base for comparison. The estimated probability of 11 out of the 13 fixed effects for Africa versus Europe were significant (see Table 31). African parents' attitudes towards pregnant girls differed from those of Europeans and North Americans on the following items: 1) a pregnant teenage girl a deviant, 2) teenage pregnancy makes parents devastated, 3) teenage pregnancy means that parents have failed, 4) parents should stop her schooling, 5) parents should give her away in early marriage, 6) parents should abort pregnancy, and 7) parents should send her to a maternity group home for pregnant teenage girls for care, and

The attitudes of African parents did not differ from those of Europeans and North America on the following items: 1) a pregnant teenage girl do not have future, and 2) teenage pregnancy results in stigmatization. However, while Europeans differ from Africans on 1) teenage pregnancy is the result of Western education, 2) parents should drive girl away from the family in anger, 3) parents should support girl to have the baby, and 4) parents should force the man/boy responsible for the pregnancy to accept responsibility of her care, North America didn't differ (Table 31).

Table 31 Teenage pregnancy replication region comparisons for the full sample, and heterogeneity test

Item of test	Africa v. Europe		Africa v. Europe		<i>Predicted</i> %	<i>Predicted</i> %	<i>Predicted</i> %	Random effects			
	Estimate	95%CI Lower, Upper	Estimate	95%CI Lower, Upper				<i>sd</i> _{Site(Overall)}	<i>sd</i> _{Site(Europe)}	<i>sd</i> _{Site(Africa)}	<i>sd</i> _{Site(North America)}
<i>Item 1: A pregnant teenage girl is a deviant</i>	0.39*	0.18, 0.87	0.37*	0.15, 0.90	0.149	0.306	0.140	1.09	1.71	0.44	0.14
<i>Item 2: A pregnant teenage girl do not have future</i>	0.99	0.62, 1.59	0.59	0.59, 0.59	0.057	0.057	0.035	0.47	0.42	0.76	0.31
<i>Item 3: Teenage pregnancy makes parents devastated</i>	1.14***	0.08, 0.24	0.19***	0.19, 0.19	0.355	0.796	0.437	0.72	1.07	0.39	0.17
<i>Item 4: Teenage pregnancy means that parents have failed</i>	0.60*	0.41, 0.89	0.26***	0.16, 0.42	0.105	0.163	0.049	0.45	0.59	0.20	0.38
<i>Item 5: Teenage pregnancy results in stigmatization</i>	1.15	0.73, 1.81	1.07	0.64, 1.79	0.783	0.759	0.771	0.59	0.56	0.61	0.59
<i>Item 6: Teenage pregnancy is the result of western education</i>	0.39***	0.28, 0.55	0.80	0.55, 1.18	0.088	0.199	0.166	0.39	0.28	0.57	0.28
<i>Item 7: Parents should drive girl away from the family in anger</i>	0.35**	0.17, 0.72	0.49	0.23, 1.08	0.010	0.029	0.015	0.68	0.53	0.86	0.67
<i>Item 8: Parents should stop her schooling</i>	0.26***	0.15, 0.45	0.32***	0.17, 0.59	0.019	0.069	0.023	0.43	0.90	0.35	0.34
<i>Item 9: Parents should give her out in early marriage</i>	0.29***	0.14, 0.59	0.35*	0.16, 0.79	0.027	0.088	0.033	0.89	1.29	0.44	0.41
<i>Item 10: Parents should abort pregnancy</i>	2.22*	1.19, 4.13	2.04*	1.01, 4.13	0.190	0.096	0.178	0.80	0.87	0.00	0.50

Increasing the generalizability of psychology by including Africans

<i>Item 11: Parents should send her to a maternity group home for pregnant teenage girls for care</i>	0.29***	0.17, 0.49	0.38**	0.38, 0.38	0.137	0.354	0.174	0.69	0.93	0.59	0.40
<i>Item 12: Parents should support girl to have the baby</i>	0.49*	0.27, 0.89	0.81	0.41, 1.60	0.772	0.873	0.847	0.79	1.03	0.59	0.49
<i>Item 13: Parents should force the man/boy responsible for the pregnancy to accept responsibility of her care</i>	0.47***	0.32, 1.60	0.89	0.57, 1.41	0.371	0.553	0.525	0.52	0.46	0.73	0.31

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

A summary table showing the estimates of the fixed effects for the subgroup between regions for Africa versus Europe for the 13 items on attitude towards teenage pregnancy and their %CIs, the predicted probability per region, and by-site standard deviations. Also, the by-site standard deviations for Africa and Europe subgroups for the within region heterogeneity test.

Appendix VII

1. Condom use negotiation study results for the full sample

We hypothesized that male participants had 1) more favorable attitudes towards condom use, 2) greater feelings of control over condom use, and 3) higher self-reported condom-protected sexual behaviors, than their female counterparts. We present our findings in Table 32 for the focal effect estimates of gender on attitudes towards condom use, control over condom use, and condom-protected sexual behavior, and its fixed effect estimates comparing Africa versus Europe and Africa versus North America, and their CIs for the full sample.

1.1 Main effect of gender

When we used our model (see Table 9) to test the focal effects for the full sample. We observed that female participants reported a marginally more positive attitude towards condom use than male participants ($b_{gender} = .15$; $t_{gender(29.36)} = 2.31$, $sd(b_{gender})_{gender} = .26$, $p = .028$; see Figure 9). However, male participants reported slightly greater control over condom use ($b_{gender} = -.23$; $t_{gender(27.72)} = -3.21$, $sd(b_{gender})_{gender} = .10$, $p = .003$; see Figure 10) and higher condom-protected sexual behavior ($b_{gender} = -.49$; $t_{gender(33.07)} = -9.89$, $df = 33.07$, $sd(b_{gender})_{gender} = .14$, $p < .05$; see Figure 11) than their female counterparts. Although women hold slightly more positive towards condom use than men, men are marginally more likely to express higher control over condom use, and self-report more condom-protected sexual behavior than women.

1.2 Gender by region comparison for the entire participants

To test whether the focal effects differ by region, we fitted Africa_vs_Europe and Africa_vs_North America as fixed slopes in our model to test the main and interaction effects of gender on attitude, control, and condom-protected sexual behavior across the entire participants and for the region comparisons using Africa as base for comparison. We found that the main effects of gender across the entire participants and its interaction effects for the region comparisons were significant for all the variables except for attitude towards condom use for Africa versus North America (see Table 32).

Our heterogeneity test for the full sample showed that while by-site estimates' standard deviations for main effects of gender on attitudes towards condom use were larger than focus effect estimate for this variable, the by-site estimates' standard deviations for the main effects of gender on control over condom use and self-reported condom-protected sexual behavior were smaller than the focus effect estimates for these variable outcomes (see Table 32). Hence, across regions and within the regions, we observed variability for attitude towards condom use and no variability for control over condom use and self-reported condom-protected sexual behavior for the full sample.

Table 32 Condom use negotiation results for the overall analyses and the replication test

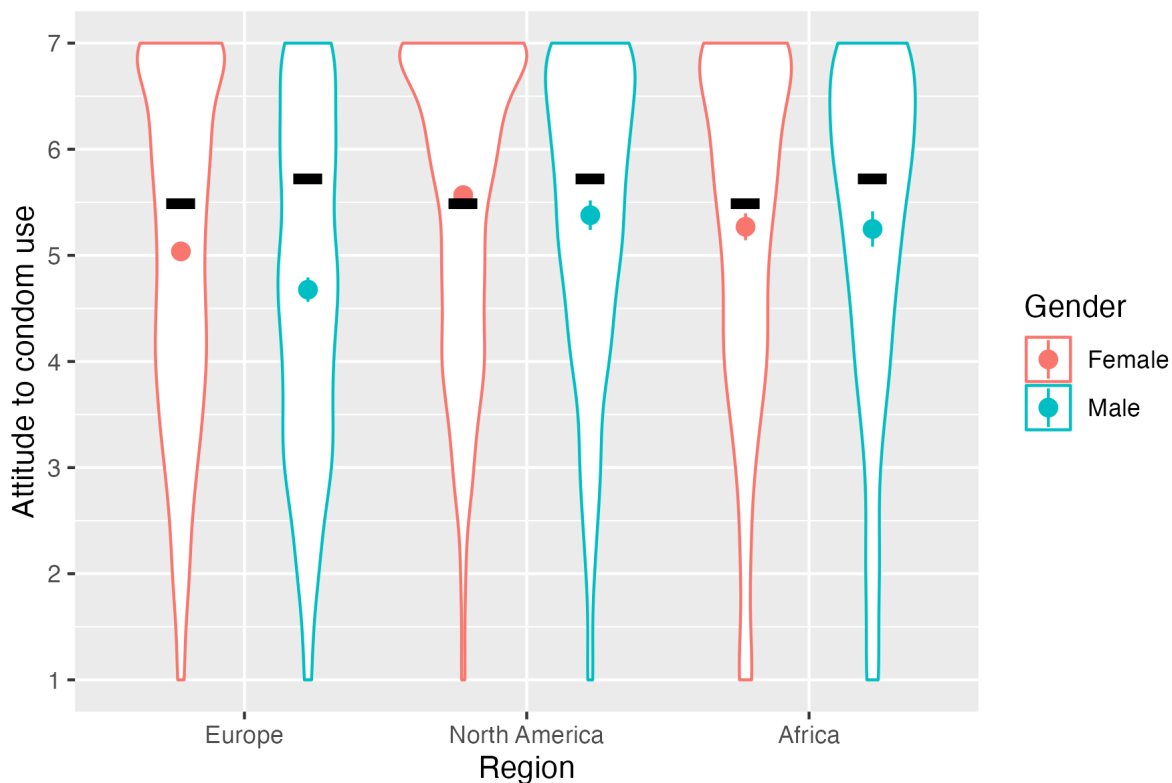
Outcome variable	Fixed effects		95%CI		SE	t-value	df	p-value	Random effects				Replication test		
	Parameter	Original estimate	Replication estimate	Lower Bound					Upper Bound	sd _{Site(O} verall)	sd _{Site(E} urope)	sd _{Site(Af} rica)	sd _{Site(Nort} h America)	Preset <i>a</i>	Replication decision
Attitude	<i>b_{gender}</i>	-0.23	0.15	0.017	0.29	0.07	2.31	29.36	.028	0.26	0.63	0.41	0.29	0.0167	Non replication
	<i>b_{gender*africa_v_eu}</i>	-	0.38	0.06	0.69	0.16	2.39	34.64	.022	0.26	-	-	-	-	
	<i>b_{gender*africa_v_us}</i>	-	0.14	-0.22	0.49	0.17	0.79	32.53	.438	0.26	-	-	-	-	
Control	<i>b_{gender}</i>	-0.34	-0.23	-0.05	0.04	0.04	-3.21	27.72	.003	0.10	0.08	0.12	0.12	0.0167	Replication
	<i>b_{gender*africa_v_eu}</i>	-	0.49	0.27	0.70	0.11	4.59	34.64	<i>p</i> < .05	0.10	-	-	-	-	
	<i>b_{gender*africa_v_us}</i>	-	0.35	0.12	0.59	0.12	3.04	34.21	.005	0.10	-	-	-	-	
Protected sexual behavior	<i>b_{gender}</i>	-0.74	-0.49	-0.52	-0.39	0.05	-9.89	33.07	<i>p</i> < .05	0.14	0.44	0.32	0.13	0.0167	Replication
	<i>b_{gender*africa_v_eu}</i>	-	0.46	0.22	0.70	0.12	3.83	41.27	<i>p</i> < .05	0.14	-	-	-	-	

$b_{gender*africa_v_us}$	-	0.28	0.01	0.54	0.13	2.12	38.89	.041	0.14	-	-	-	-
----------------------------	---	------	------	------	------	------	-------	------	------	---	---	---	---

Note. The table provides a summary of the original fixed effect estimate derived from the male and female mean differences for the three outcome variables. The replication fixed effects estimates of gender on attitudes towards to condom use (“attitude”), perceived control over condom use (“control”), and self-reported condom-protected sexual behavior (“protected sexual behavior”) for the full sample, and for Africa versus Europe, and for Africa versus North America, and as well as by-site variability for these effects, and their CIs.

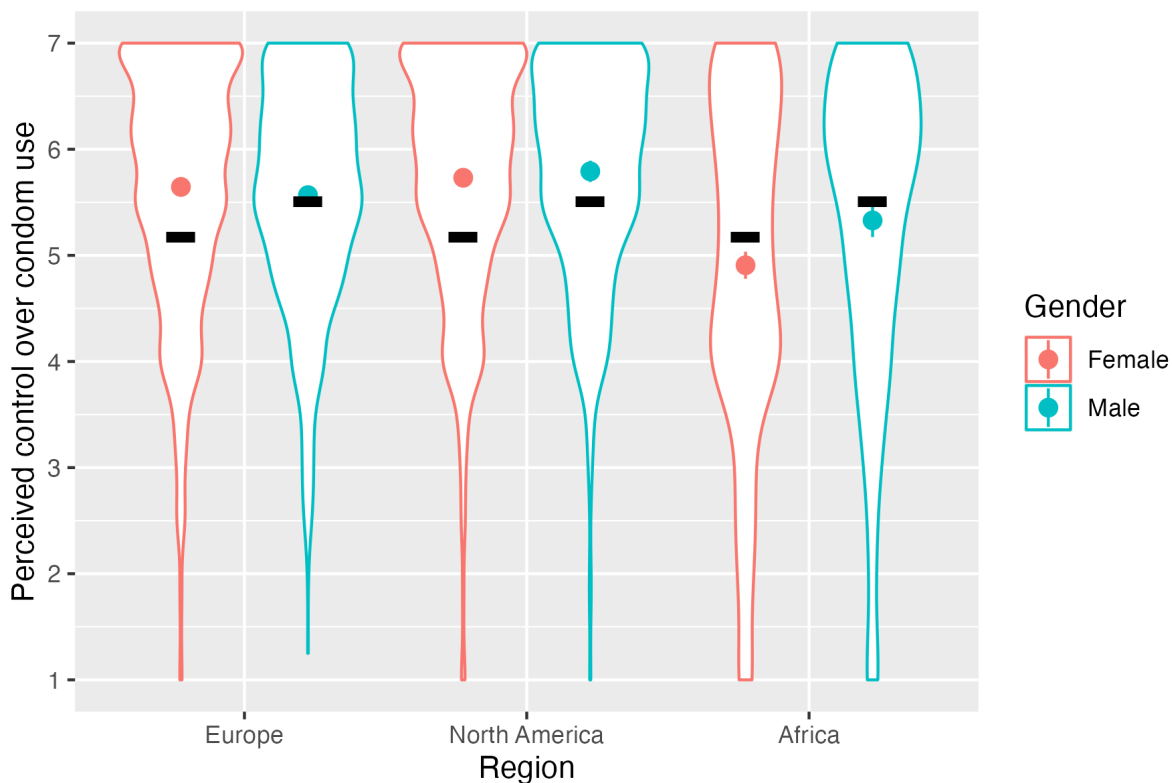
Given the obtained p -values for the main effects for gender on attitude towards condom use ($b = .15$, p -value = .028), control over condom use ($b = -.23$, p -value = .003), and self-reported condom-protected sexual behavior ($b = -.49$, p -value = .00000000002), we compared the replication p -values to our preset alpha (0.0167), and found a replication for the effect of gender on perceived control over condom use and self-reported condom-protected sexual behavior.

Figure 9 Attitude to condom use by gender per region for the full sample



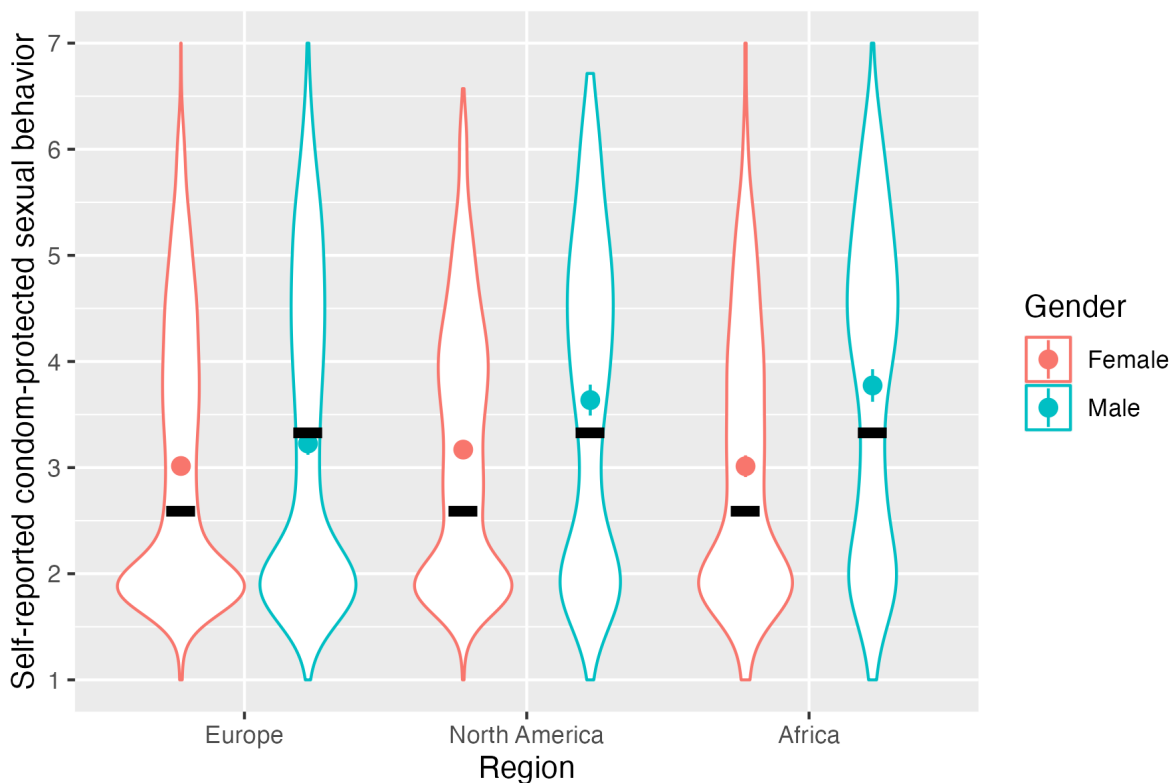
Note. The average scores between male and female on attitude towards condom use for the North Americans and Africans were marginal but it was considerable for Europeans. In each violin are the original study male and female students' attitude towards condom use average scores (black crossbars). The original study's attitude towards condom use means for African adolescent both male and female college students are larger than our replication mean scores for male and female participants except for the North American female sample.

Figure 10 Perceived control over condom use by gender per region for the full sample



Note. The average scores on perceived control over condom use were marginal between male (blue dots) and female (red dots) for both Europeans and North Americans, but considerable between male and female Africans. Also, African male and female participants mean scores are lesser compared to their European and North American counterparts. In each violin are the original study of male and female students’ perceived control over condom use average scores (black crossbars). The original mean scores are larger than our replication mean scores for Africans but lesser than Europeans and North Americans male and female average scores.

Figure 11 Self-reported condom-protected sex behavior by gender per region the full sample



Note. Across all the regions of Europe, North America, and Africa, male average scores on self-reported condom-protected sexual behavior (blue dots) were higher than for their female counterparts (red dots). Also, in each violin plot for all regions are the original study male and female students' self-reported condom-protected sexual behavior average scores (black crossbars). The original study self-reported condom-protected sexual behavior means for adolescent male college students are lower than our replication mean scores for male and female students across all regions except for the European male category.

Appendix VIII

1. Extramarital sexual relationship study results

As earlier tested for the heterosexual women, we tested two hypotheses however, for the full sample. We tested whether participants were more willing to forgive 1) a husband with a low degree of emotional involvement in an extramarital affair compared to a husband with a high degree emotional involvement in an extramarital affair, and 2) when a husband expressed high regret compared to when a husband expressed low regret for engaging in an extramarital sexual relationship. We present the extramarital sexual relationship study results in Table 32 for the full sample main effect estimates for emotional involvement and regret expressed, and the two-way interaction effect estimates comparing Africa versus Europe and Africa versus North America, and their CIs.²⁵

1.1 Main effect of emotional involvement and regret expressed on forgiveness for the full sample

We analyzed combined forgiveness ratings for the entire participants using our linear mixed effects model. We fitted dummy-coded indicator variables for the comparisons between Africa and either Europe or North America as fixed slopes in our model (see Table 9) using Africa as base for comparison to test the main effects of emotional involvement and regret expressed, and their interactions with the region comparisons. Although, participants reported low willingness to forgive a cheating partner regardless whether the regret expressed or

²⁵ We did not present 1) main effects of commitment to marriage, 2) the two-way interactions between these three factors and 3) the three-way and 4) the four-way interactions for these variables. The full results are posted at <https://osf.io/6cbmw>. We trimmed the random effects for the interaction terms because the model was not identifiable with so many random effects, and we decided that the interaction terms were not of primary importance given the study's focus on main effects.

emotional involvement was low or high (see Figure X), they are more likely to forgive a cheating husband 1) that expressed high regret compared to one that expressed low regret ($b_{regret} = 1.27, p < .05$) and 2) less emotionally involved than one that is highly emotionally involved in an extramarital affair ($b_{emotion} = -.28, p < .05$, see Table 20). Participants reported the highest average forgiveness ratings for low emotional involvement*high regret expressed condition ($M_{low\ emotion*high\ regret} = 2.18, sd = 2.57$). While they reported the lowest average ratings for high emotional involvement and low regret condition ($M_{high\ emotion*low\ regret} = 0.77, sd = 1.68$). These results were consistent when we compared Africa versus Europe and Africa versus North America. African participants ($M_{Africa} = 2.95, sd = 3.07$) were almost thrice more willing to forgive a cheating partner than their European ($M_{Europe} = 1.06, sd = 1.84$) and North American ($M_{North\ America} = 1.02, sd = 1.69$) counterparts. European participants reported the lowest average rating (Av. rating = 0.47, $sd = 1.19$) when the regret expressed was low and emotional involvement was high while African participants reported the highest average forgiveness rating (Av. rating = 4.05, $sd = 3.18$) when the regret expressed was high and emotional involvement was low (see Figure 12).

Our heterogeneity test showed that by-site estimates' standard deviations for effect of emotional involvement and regret expressed on willingness to forgive were smaller than focus effect estimates for this variable, indicating little variability across regions and within the regions. Such as the subgroup, because these factors were nested in participants, and participants' responses to these scenarios should differ, tested the by-participant variability among participants and within the regions. We observed that by-participant standard deviation estimates for emotional involvement and for regret expressed (and again except for North American participants) were larger than the main effect estimates for the full sample, for Africans, for Europeans, and for North Americans. Hence, we found variability across the participants as well as within regions (see Table 33).

Table 33 Extramarital sexual relationship focal effect results for the entire participants

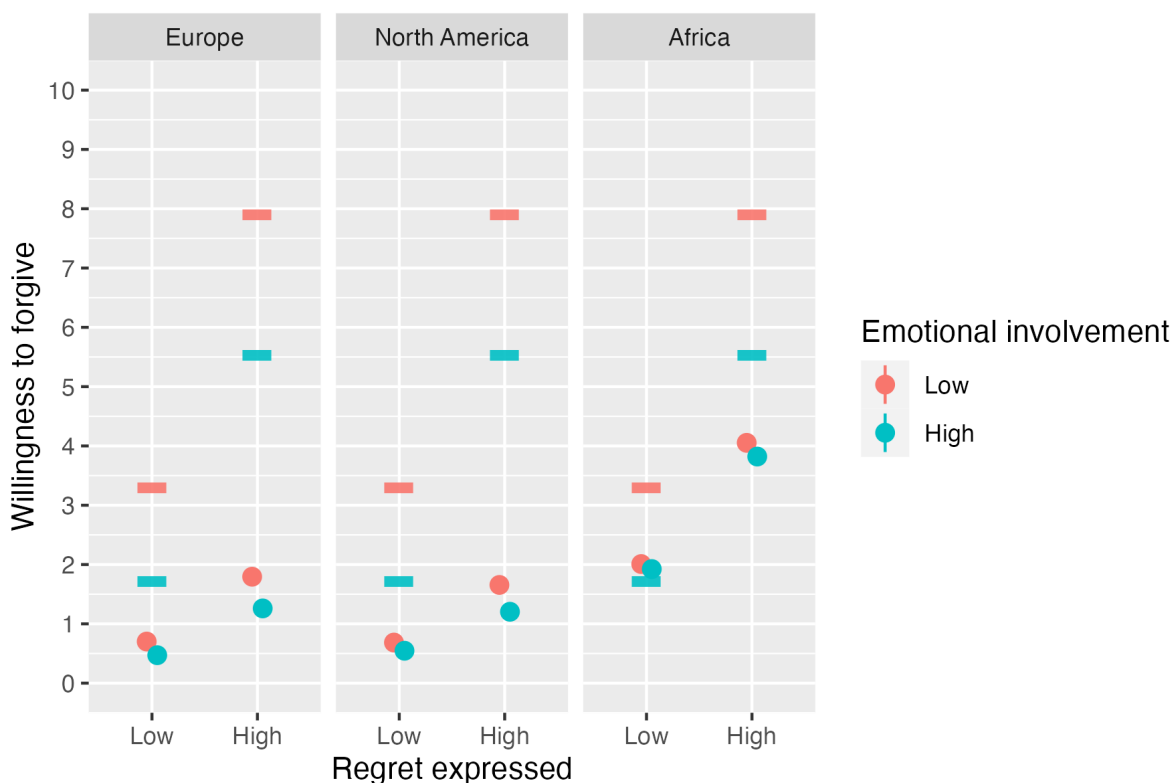
Fixed effects		Random effects														Replication test	
Parameter	Original estimate	Replication estimate	95%CI		<i>t</i> -value	<i>df</i>	<i>p</i> -value	Overall		Europe	Africa	North America		Preset <i>a</i>	Replication decision		
			Lower Bound	Upper Bound				<i>sd</i> _{ID}	<i>sd</i> _{Site}	<i>sd</i> _{ID}	<i>sd</i> _{Site}	<i>sd</i> _{ID}	<i>sd</i> _{Site}				
<i>b</i> _{emotion}	-1.97	-.28	-0.34	-0.22	-10.34	24.99	.0000000002	0.78	0.14	0.84	0.18	0.63	0.10	0.63	0.00	.0005	Replication
<i>b</i> _{regret}	4.21	1.27	1.11	1.42	17.01	36.51	.00000000000000002	1.51	0.44	1.36	0.28	2.25	0.67	1.16	0.17	.0005	Replication
<i>b</i> _{emotion*africa_v_eu}	-	-0.25	-0.38	-0.12	-3.91	28.43	.0005	0.78	0.14	-	-	-	-	-	-	-	-
<i>b</i> _{regret*africa_v_eu}	-	-1.19	-1.54	-0.85	-7.04	39.34	.000000002	1.51	0.44	-	-	-	-	-	-	-	-
<i>b</i> _{emotion*africa_v_us}	-	-0.15	-0.29	-0.00	-2.07	26.10	.049	0.78	0.14	-	-	-	-	-	-	-	-
<i>b</i> _{regret*africa_v_us}	-	-1.29	-1.69	-0.91	-6.72	37.38	.000000006	1.51	0.44	-	-	-	-	-	-	-	-

Note. Dependent variable - Willingness to forgive

A summary table showing the original and our replication main effect estimates of emotional involvement and regret expressed on willingness to forgive for the full sample. The two-way interaction effect estimates for emotional involvement by Africa_vs_Europe, for emotional involvement by Africa_vs_North America, for regret expressed by Africa_vs_Europe, and regret expressed by Africa_vs_North America, and their CIs, *t*-values, *df*, and *p*-values. We present the by-person and by-site standard deviations for these focal effects for the overall participants and for participants in each region. Lastly, possible replication interpretations of our results given *p*-values we obtain for the main effects for emotional

involvement ($b_{emotion} = -.28, p\text{-value} = .0000000002$) and regret expressed ($b_{regret} = 1.27, p\text{-value} = .000000000000000002$) for the entire participants compared to the original study preset alpha (.0005).

Figure 12 Willingness to forgive mean ratings as an operation of regret expressed and emotional involvement and for region comparison for full sample



Note. Willingness to forgive average ratings on the y-axis. Three panels correspond to the regions: Europe, North America, and Africa. For each panel we have regret expressed on the x-axis for low and high regret expressed conditions. In each panel, we have red and blue cross bars and dots indicating the combination of low versus high conditions for regret expressed and emotional involvement for the original study and replication respectively. The red cross bars/dots indicate low emotional involvement and the blue cross bars/dots high emotional involvement conditions. In the replication study, across the conditions and regions, when the regret expressed is high and the emotional involvement is low, participants are more likely to forgive a cheating husband compared to when the regret expressed is low and the emotional involvement is high. Our replication average forgiveness ratings are much lower than reported in the original study²⁶.

²⁶ We extracted the original forgiveness means from the graph presented in Vera Cruz's (2018, Figure 1) using WebPlotDigitizer (Rohatgi, 2024; <https://automeris.io/WebPlotDigitizer.html>).