

The Effects of Visual and Design Features on the Perception of Correlation in Scatterplots

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Science and Engineering

2024

Gabriel Strain
Department of Computer Science

Contents

Contents	2
List of figures	6
List of tables	7
List of publications	8
Abstract	9
Lay abstract	10
Declaration of originality	11
Copyright statement	12
Acknowledgements	13
1 Introduction	14
1.1 Research Motivation	14
1.2 Contributions	14
1.3 Included Publications	14
1.4 Overview of Thesis	14
2 Literature Review	15
2.1 Data Visualisation: A Brief History	15
2.2 Measuring Relatedness	15
2.3 Conceptions of Correlation	15
2.4 Visualising Correlation	15
2.4.1 History	15
2.4.2 Present Landscape	15
2.4.3 Scatterplots	15
2.5 Correlation Perception	15
2.6 Correlation Cognition	15
2.7 Underestimation: What's Really Going On?	15
2.8 Underestimation: Potential Consequences	15
2.9 Data Visualisation Literacy	15
2.10 Objectives and Contributions	15

3	General Methodology	16
3.1	Introduction	16
3.2	Experimental Methods	16
3.2.1	Experimental Design	16
3.2.2	Tools for Testing	17
3.2.3	Recruitment & Participants	17
3.2.4	Creating Stimuli	18
3.3	Analytical Methods	18
3.3.1	Linear Mixed-Effects Models	18
3.3.2	Advantages Over Aggregate-Level Statistical Tests	20
3.3.3	Model Construction	20
3.3.4	Effects Sizes	20
3.3.5	Reporting Analyses	20
3.4	Computational Methods	20
3.4.1	Executable Reporting	20
3.4.2	Containerised Environments	20
3.5	Reproducibility In This Thesis	20
3.5.1	Sharing Data and Code	20
3.5.2	Executable Papers and Docker Containers	20
3.5.3	Experimental Resources	20
3.6	Conclusion	21
4	Adjusting the Opacities of Scatterplot Points Can Affect Correlation Estimates	22
4.1	Abstract	23
4.2	Preface: Learning From an Early Pilot Study	23
4.3	Introduction	23
4.3.1	Overview	23
4.4	Related Work	23
4.4.1	Transparency, Contrast, Opacity, and Formal Definitions	23
4.4.2	Effects of Point Opacity on Correlation Estimation	23
4.5	Experiment 1: Uniform Opacity Adjustments	23
4.5.1	Introduction	23
4.5.2	Methods	23
4.5.3	Analysis	23
4.5.4	Discussion	23
4.6	Experiment 2: Spatially-Dependent Opacity Adjustments	23
4.6.1	Introduction	23
4.6.2	Methods	23
4.6.3	Analysis	23
4.6.4	Discussion	23
4.7	General Discussion	23

5	Adjusting the Sizes of Scatterplot Points Can Correct for a Historic Correlation Underestimation Bias	24
5.1	Abstract	24
5.2	Overview	24
5.3	Related Work	24
5.3.1	Size and Perception	24
5.3.2	Scatterplot Point Size and Correlation Perception	24
5.4	Experiment: Adjusting Point Size to Facilitate More Accurate Correlation Perception in Scatterplots	24
5.4.1	Introduction	24
5.4.2	Methods	24
5.4.3	Analysis	24
5.4.4	Discussion	24
5.5	General Discussion	24
6	Interactions of Opacity and Size Adjustments	25
6.1	Abstract	25
6.2	Overview	25
6.3	Related Work	25
6.3.1	Size and Opacity	25
6.4	Experiment: Adjusting Point Size and Opacity Together	25
6.4.1	Introduction	25
6.4.2	Methods	25
6.4.3	Analysis	25
6.4.4	Discussion	25
6.5	General Discussion	25
7	Visual Features Affecting Perceptual Estimates Also Affect Beliefs About Correlations	26
7.1	Abstract	27
7.2	Overview	27
7.3	Related Work	27
7.3.1	From Perception to Cognition	27
7.3.2	From Cognition to Belief	27
7.4	Pre-Study: Investigating Beliefs About Relatedness Statements	27
7.4.1	Introduction	27
7.4.2	Methods	27
7.4.3	Analysis	27
7.4.4	Discussion	27
7.5	Experiment: Potential for Belief Change Using Atypical Scatterplots	27
7.5.1	Introduction	27
7.5.2	Methods	27
7.5.3	Analysis	27
7.5.4	Discussion	27

7.6 General Discussion	27
8 Conclusion	28
8.1 Main Findings	28
8.2 Relationship to Prior Work	28
8.3 Reproducibility	28
8.4 Contributions	28
8.5 Implications	28
8.5.1 For Design	28
8.5.2 For Society	28
8.6 Limitations	28
8.7 Future Directions	28
8.8 Closing Remarks	28
References	29
Appendices	31
A First appendix	32

Word count: 1000

List of figures

3.1	An example of the slider participants used to estimate correlation in experiments 1-4. . .	17
3.2	The basic design of scatterplots in experiments 1 to 4.	19

List of tables

List of publications

Publications go here.

Abstract

put abstract here

Lay abstract

This is lay abstract text.

Declaration of originality

I hereby confirm that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

- i The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made *only* in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.library.manchester.ac.uk/about/regulations/>) and in The University’s policy on Presentation of Theses.

Acknowledgements

Acknowledgements go here.

Chapter 1

Introduction

1.1 Research Motivation

1.2 Contributions

1.3 Included Publications

The research described in chapters 4, 5, 6, and 7 in this thesis is adapted from earlier publications, the last of which under review as of writing. To avoid repetition, information and discussion that would be repeated has been consolidated into the literature review and general methodology chapters. *Gabriel Strain* is the primary author of all included papers.

- *The Effects of Contrast on Correlation Perception in Scatterplots* [15] is reproduced in Chapter 4. Sections 4.5.2, 4.6.2, 4.5.3, 4.6.3, 4.5.4, 4.6.4, and 4.7 contain minimally altered parts of the published article.
- *Adjusting Point Size to Facilitate More Accurate Correlation Perception in Scatterplots* [14] is reproduced in Chapter 5. Sections 5.4.2, 5.4.3, 5.4.4, and 5.5 contain minimally altered parts of the published article.
- *Effects of Point Size and Opacity Adjustments in Scatterplots* [16] is reproduced in Chapter 6. Sections 6.4.2, 6.4.3, 6.4.4, and 6.5 contain minimally altered parts of the published article.
- *Effects of Alternative Scatterplot Designs on Belief (under review)* is reproduced in Chapter 7. Sections 7.4, 7.5.2, 7.5.3, 7.5.4, and 7.6 contain minimally altered parts of the published article.

1.4 Overview of Thesis

Chapter 2

Literature Review

2.1 Data Visualisation: A Brief History

2.2 Measuring Relatedness

2.3 Conceptions of Correlation

2.4 Visualising Correlation

2.4.1 History

2.4.2 Present Landscape

2.4.3 Scatterplots

2.5 Correlation Perception

2.6 Correlation Cognition

2.7 Underestimation: What's Really Going On?

2.8 Underestimation: Potential Consequences

2.9 Data Visualisation Literacy

2.10 Objectives and Contributions

Chapter 3

General Methodology

3.1 Introduction

In this chapter we describe our research methodologies. Chapters 4, 5, and 6 share most aspects of experimental method, while the experiment described in chapter 7 differs substantially. Throughout this chapter, the reader should assume that we are referring to the entire body of experimental work this thesis describes. Methods that differ regarding the final experiment in chapter 7 are detailed along the way. In this chapter, we discuss our experimental designs, the tools we use to build and run our experiments, our approach to statistical analyses, and the computational methods and practices we employed particularly with regards to reproducibility and open science.

3.2 Experimental Methods

It is important to acknowledge that the way in which we conduct experiments influences what we find and the conclusions that we may draw from those findings. The decisions that lead us to designing experiments in certain ways must be based not only on theory, but also on the practical constraints imposed by external factors on the research team. Concerns such as time, convenience, and cost must be addressed, and a compromise between research that is *valuable* and research that is *doable* must be reached. We focused on pragmatism and impact throughout the course of this research project; happily, the research journey we embarked on resulted in methodologies that satisfied both principles. It is for this reason that we consider the framework we present to be a key contribution of this thesis.

3.2.1 Experimental Design

All but our final experiment utilised within-participants designs. Each participant saw all experimental stimuli and provided a judgement of correlation using a sliding scale between 0 and 1 (see Figure 3.1). Experiments 1 to 3 featured a single experimental factor of design, all with 4 levels corresponding to scatterplots with different design features. Experiment 4 employed a factorial 2×2 design. Experiment 5 is a departure from the shared experimental paradigm of the previous experiments, and features a 1 factor, 2 level between-participants design.

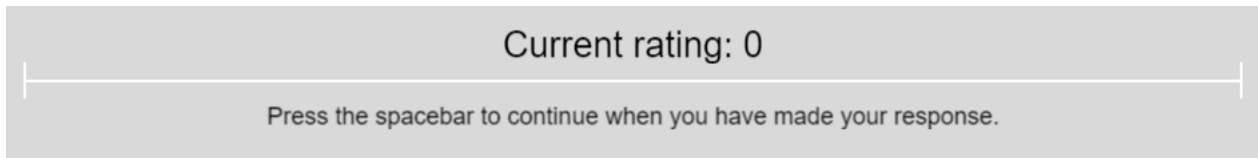


Figure 3.1. An example of the slider participants used to estimate correlation in experiments 1-4.

3.2.2 Tools for Testing

Whatever the design of our experiments, software plays a crucial role in allowing us to carry them out. Fortunately, at the time of writing, there is a wealth of tools available to facilitate the testing of visualisations both in traditional lab-based tests and in online experiments. As we adhere to the principles of open and reproducible research [3], we discount closed-source software, such as Gorilla [1] or E-prime [7], as these rely on paid licenses and do not allow us to share code with future researchers. We settled on using PsychoPy [10] due to its open-source status, flexibility regarding graphical and code-based experimental design, and high level of timings accuracy [4]. Using such an open-source tool not only facilitated our own learning with regard to experiment building, but also enables to contribute further examples of visualisation studies by hosting the resulting experiments online for use and modification by future researchers.

We elected to pursue online testing throughout this thesis. Doing so is much quicker than carrying out in-person lab-based testing, meaning we can collect data from a much larger number of participants. This reduces the chances of detecting false positives during analysis and ensures adequate levels of power despite the potential for small effects sizes. Online testing also affords us access to diverse groups of participants across our populations of interest, especially when compared to the relatively homogeneous student populations usually accessed by doctoral researchers. Research has identified online experimentation as producing reliable results that closely match those found in traditional lab-based experiments [2, 8, 11], especially with large sample sizes. Due to its integration with PsychoPy, we chose to use Pavlovia (pavlovia.org) to host all the experiments described in this thesis. Section 3.5.3 contains links to all experiments publicly hosted on Pavlovia's GitLab instance.

3.2.3 Recruitment & Participants

Recruitment of participants online is possible through a range of service providers, each with advantages and disadvantages. Research evaluating a number of these providers recently found that Prolific [12] and CloudResearch provide the highest quality data for the lowest cost [6]; we elected to use the former due to familiarity with the system. Despite these findings, there has also been evidence of low data quality and skewed demographics affecting even high quality platforms tailored towards academic research. On the 24th of July, 2021, the Prolific.co platform went viral on social media [5], leading to a participant pool heavily skewed towards young people identifying as female. At the time, Prolific did not manually balance the participants recruited for a study. We addressed this in our pilot study (see Section 4.2) by preventing participants who joined after this data from participating, in addition to manually requesting a 1:1 ratio of male to female participants. The demographic issues settled quickly, however we maintained our screened 1:1 ratio for the remainder of the experiments.

The first experiment we conducted was a pilot study (see Section 4.2 for full details) investigating a very early iteration of the point opacity manipulation in combination with exploratory work around plot size and correlation estimation. At the time, the author was relatively naive to the intricacies of recruiting research participants online, and thus experienced issues with regards to participant engagement. Each experiment, including the pilot, included attention check questions in which participants were instructed to ignore the stimulus and provide a specific answer. We stated in the advert for each experiment that failure of more than 2 attention check items would result in a submission being rejected. This pilot study suffered from a rejection rate of 57.5%, indicating that we were experiencing low levels of participant engagement. For our following studies, we therefore followed published guidelines [9] to address these issues; specifically, we required that participants:

- Had previously completed at least 100 studies on Prolific.
- Had an acceptance rate of at least 99% for those studies.¹

Following implementation of these pre-screen criteria, the rejection rate for our next experiment fell to ~5%. Rejection rates were similar for the remainder of our experiments. Exact numbers of accepted and rejected participants can be found in the **Participants** sections of each experiment.

3.2.4 Creating Stimuli

All our stimuli were created using `ggplot2` in R. Specific versions are cited separately with regard to the specific visualisations produced for each experiment. We followed identical principles regarding data visualisation design for each experiment bar the last, which is discussed *in situ*.

We designed with the intention of isolating and addressing a perceptual effect; the underestimation of correlation in positively correlated scatterplots. For this reason, we sought to remove the potential for other design factors to have effects on correlation estimation. To this end, we removed most of the conventionally present visual features of scatterplots, including axis labels, tick labels, grid lines, and titles. We elected to preserve the axis ticks themselves. Figure 3.2 demonstrates the basic design of the scatterplots used in experiments 1 to 4.

3.3 Analytical Methods

3.3.1 Linear Mixed-Effects Models

To investigate whether the experimental manipulations we test have actual effects on the interpretations participants provide, we must employ appropriate statistical testing. This involves taking into account the variability in responses that can be attributed to an experimental experimental against the backdrop of other variability inherent in the dataset. To accomplish this, we utilise linear mixed-effects modelling, a broadly applicable and reliable approach that is also resistant to a variety of distributional assumption violations [13].

¹this is a more strict rate than the 95% recommended by Peer et al. [9].

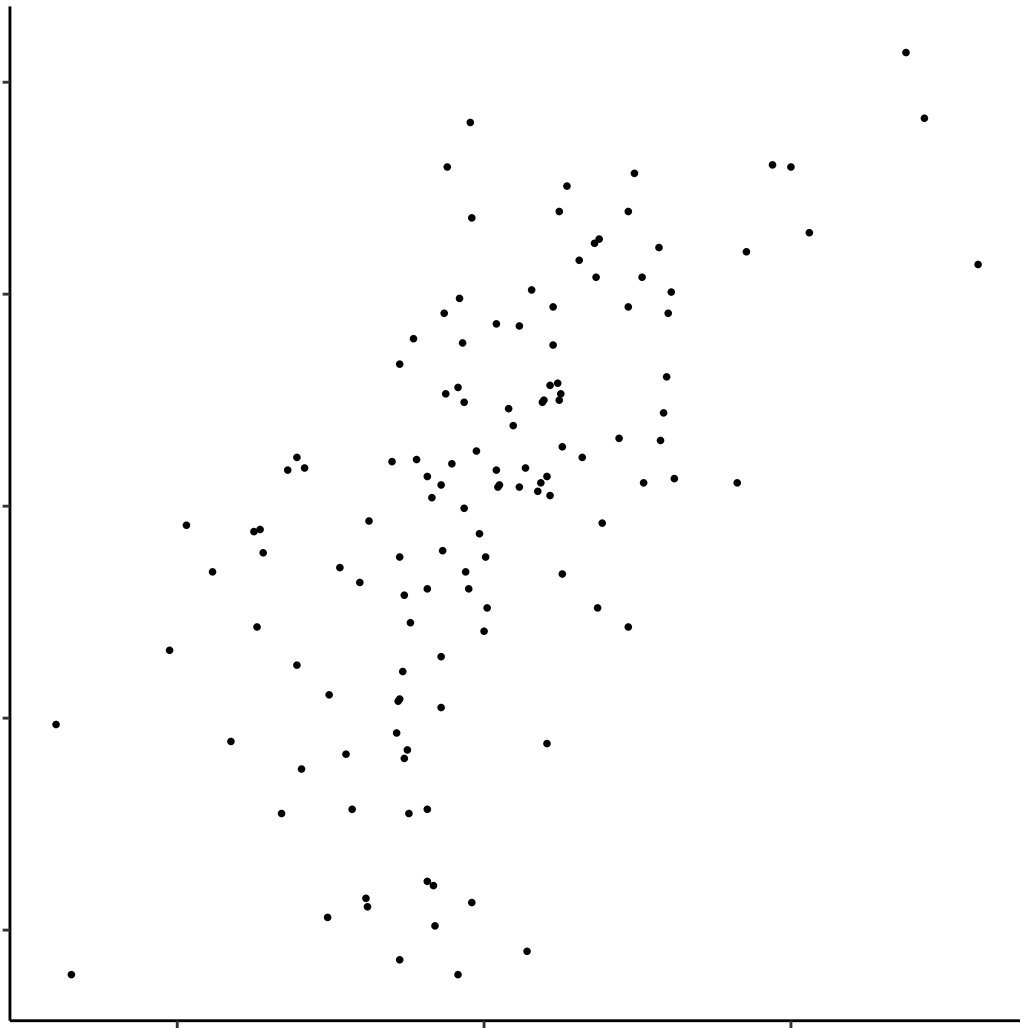


Figure 3.2. The basic design of scatterplots in experiments 1 to 4.

- fixed and random effects
- how does modelling work?

3.3.2 Advantages Over Aggregate-Level Statistical Tests

3.3.3 Model Construction

3.3.4 Effects Sizes

3.3.5 Reporting Analyses

3.4 Computational Methods

3.4.1 Executable Reporting

3.4.2 Containerised Environments

3.5 Reproducibility In This Thesis

3.5.1 Sharing Data and Code

3.5.2 Executable Papers and Docker Containers

3.5.3 Experimental Resources

Everything needed to run each experiment is included in the corresponding GitLab repository. Links to these repositories are also provided in the sections concerning each experiment.

Chapter 4

Experiment 1: https://gitlab.pavlovia.org/Strain/exp_uniform_adjustments

Experiment 2: https://gitlab.pavlovia.org/Strain/exp_spatially_dependent

Chapter 5

Experiment 3: https://gitlab.pavlovia.org/Strain/exp_size_only

Chapter 6

Experiment 4: https://gitlab.pavlovia.org/Strain/size_and_opacity_additive_exp

Chapter 7

Experiment 5 Pre-Study: https://gitlab.pavlovia.org/Strain/beliefs_scatterplots_pretest

Experiment 5 Main Study (Group A): https://gitlab.pavlovia.org/Strain/atypical_scatterplots_main_a

Experiment 5 Main Study (Group B): https://gitlab.pavlovia.org/Strain/atypical_scatterplots_main_t

3.6 Conclusion

Chapter 4

Adjusting the Opacities of Scatterplot Points Can Affect Correlation Estimates

4.1 Abstract

4.2 Preface: Learning From an Early Pilot Study

4.3 Introduction

4.3.1 Overview

4.4 Related Work

4.4.1 Transparency, Contrast, Opacity, and Formal Definitions

4.4.2 Effects of Point Opacity on Correlation Estimation

4.5 Experiment 1: Uniform Opacity Adjustments

4.5.1 Introduction

4.5.2 Methods

4.5.3 Analysis

4.5.4 Discussion

4.6 Experiment 2: Spatially-Dependent Opacity Adjustments

4.6.1 Introduction

4.6.2 Methods

4.6.3 Analysis

Chapter 5

Adjusting the Sizes of Scatterplot Points Can Correct for a Historic Correlation Underestimation Bias

5.1 Abstract

5.2 Overview

5.3 Related Work

5.3.1 Size and Perception

5.3.2 Scatterplot Point Size and Correlation Perception

5.4 Experiment: Adjusting Point Size to Facilitate More Accurate Correlation Perception in Scatterplots

5.4.1 Introduction

5.4.2 Methods

5.4.3 Analysis

5.4.4 Discussion

5.5 General Discussion

Chapter 6

Interactions of Opacity and Size Adjustments

6.1 Abstract

6.2 Overview

6.3 Related Work

6.3.1 Size and Opacity

6.4 Experiment: Adjusting Point Size and Opacity Together

6.4.1 Introduction

6.4.2 Methods

6.4.3 Analysis

6.4.4 Discussion

6.5 General Discussion

Chapter 7

Visual Features Affecting Perceptual Estimates Also Affect Beliefs About Correlations

7.1 Abstract

7.2 Overview

7.3 Related Work

7.3.1 From Perception to Cognition

7.3.2 From Cognition to Belief

7.4 Pre-Study: Investigating Beliefs About Relatedness Statements

7.4.1 Introduction

7.4.2 Methods

7.4.3 Analysis

7.4.4 Discussion

7.5 Experiment: Potential for Belief Change Using Atypical Scatterplots

7.5.1 Introduction

7.5.2 Methods

7.5.3 Analysis

7.5.4 Discussion

Chapter 8

Conclusion

8.1 Main Findings

8.2 Relationship to Prior Work

8.3 Reproducibility

8.4 Contributions

8.5 Implications

8.5.1 For Design

8.5.2 For Society

8.6 Limitations

8.7 Future Directions

8.8 Closing Remarks

References

- [1] Alexander L. Anwyl-Irvine, Jessica Massonnié, Adam Flitton, Natasha Kirkham, and Jo K. Evershed. “Gorilla in Our Midst: An Online Behavioral Experiment Builder”. In: *Behavior Research Methods* 52.1 (Feb. 2020), pp. 388–407. ISSN: 1554-3528. DOI: 10.3758/s13428-019-01237-x. (Visited on 10/03/2024) (cited on p. 17).
- [2] Antonio A. Arechar, Simon Gächter, and Lucas Molleman. “Conducting Interactive Experiments Online”. In: *Experimental Economics* 21.1 (Mar. 2018), pp. 99–131. ISSN: 1573-6938. DOI: 10.1007/s10683-017-9527-2. (Visited on 10/04/2024) (cited on p. 17).
- [3] Paul Ayris, Isabel Bernal, Valentino Cavalli, Bertil Dorch, Jeannette Frey, Martin Hallik, Kistina Hormia-Poutanen, Ignasi Labastida i Juan, John MacColl, Agnès Ponsati Obiols, Simone Sacchi, Frank Scholze, Birgit Schmidt, Anja Smit, Adam Sofronijevic, Jadranka Stojanovski, Martin Svoboda, Giannis Tsakonas, Matthijs van Otegem, Astrid Verheusen, Andris Vilks, Wilhelm Widmark, and Wolfram Horstmann. “LIBER Open Science Roadmap”. In: (July 2018). DOI: 10.20350/digitalCSIC/15061. (Visited on 09/13/2023) (cited on p. 17).
- [4] David Bridges, Alain Pitiot, Michael R. MacAskill, and Jonathan W. Peirce. “The Timing Mega-Study: Comparing a Range of Experiment Generators, Both Lab-Based and Online”. In: *PeerJ* 8 (July 2020), e9414. ISSN: 2167-8359. DOI: 10.7717/peerj.9414. (Visited on 10/03/2024) (cited on p. 17).
- [5] Nick Charalambides. *We Recently Went Viral on TikTok - Here’s What We Learned*. <https://www.prolific.com/resources/we-recently-went-viral-on-tiktok-heres-what-we-learned>. Aug. 2021. (Visited on 10/04/2024) (cited on p. 17).
- [6] Benjamin D. Douglas, Patrick J. Ewell, and Markus Brauer. “Data Quality in Online Human-Subjects Research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA”. In: *PLOS ONE* 18.3 (Mar. 2023), e0279720. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0279720. (Visited on 10/04/2024) (cited on p. 17).
- [7] *E-Prime*. Psychology Software Tools. 2020 (cited on p. 17).
- [8] Naoyasu Hirao, Koyo Koizumi, Hanako Ikeda, and Hideki Ohira. “Reliability of Online Surveys in Investigating Perceptions and Impressions of Faces”. In: *Frontiers in Psychology* 12 (Sept. 2021). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2021.733405. (Visited on 10/04/2024) (cited on p. 17).

- [9] Eyal Peer, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer. “Data Quality of Platforms and Panels for Online Behavioral Research”. In: *Behavior Research Methods* 54.4 (Sept. 2021), pp. 1643–1662. ISSN: 1554-3528. DOI: 10.3758/s13428-021-01694-3. (Visited on 07/05/2022) (cited on p. 18).
- [10] Jonathan Peirce, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. “PsychoPy2: Experiments in Behavior Made Easy”. In: *Behavior Research Methods* 51.1 (Feb. 2019), pp. 195–203. ISSN: 1554-3528. DOI: 10.3758/s13428-018-01193-y (cited on p. 17).
- [11] Benjamin Prissé and Diego Jorrat. “Lab vs Online Experiments: No Differences”. In: *Journal of Behavioral and Experimental Economics* 100 (Oct. 2022), p. 101910. ISSN: 2214-8043. DOI: 10.1016/j.socec.2022.101910. (Visited on 10/04/2024) (cited on p. 17).
- [12] *Prolific.Co*. Prolific. 2024 (cited on p. 17).
- [13] Holger Schielzeth, Niels J. Dingemanse, Shinichi Nakagawa, David F. Westneat, Hassen Allegeue, Céline Teplitsky, Denis Réale, Ned A. Dochtermann, László Zsolt Garamszegi, and Yimeng G. Araya-Ajoy. “Robustness of Linear Mixed-Effects Models to Violations of Distributional Assumptions”. In: *Methods in Ecology and Evolution* 11.9 (2020), pp. 1141–1152. ISSN: 2041-210X. DOI: 10.1111/2041-210X.13434. (Visited on 11/29/2023) (cited on p. 18).
- [14] Gabriel Strain, Andrew J. Stewart, Paul Warren, and Caroline Jay. “Adjusting Point Size to Facilitate More Accurate Correlation Perception in Scatterplots”. In: *2023 IEEE Vis X Vision*. Melbourne, Australia: IEEE, Oct. 2023, pp. 1–5. ISBN: 9798350329841. DOI: 10.1109/VisXVision60716.2023.00006. (Visited on 02/15/2024) (cited on p. 14).
- [15] Gabriel Strain, Andrew J. Stewart, Paul Warren, and Caroline Jay. “The Effects of Contrast on Correlation Perception in Scatterplots”. In: *International Journal of Human-Computer Studies* 176 (Aug. 2023), p. 103040. ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2023.103040. (Visited on 04/11/2023) (cited on p. 14).
- [16] Gabriel Strain, Andrew J. Stewart, Paul A. Warren, and Caroline Jay. “Effects of Point Size and Opacity Adjustments in Scatterplots”. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI ’24. New York, NY, USA: Association for Computing Machinery, May 2024, pp. 1–13. ISBN: 9798400703300. DOI: 10.1145/3613904.3642127. (Visited on 05/29/2024) (cited on p. 14).

Appendices

Appendix A

First appendix

A.1 Section in Appendix