# Bayesian Statistics: A Practical Introduction for Computer Graphics

MIRINDA WHITAKER, University of Utah, USA
SARAH CREEM-REGEHR, University of Utah, USA
BOBBY BODENHEIMER, Vanderbilt University, USA

This course introduces Bayesian statistics, which is a statistical framework that can be used to analyze user studies in computer graphics and related fields. Aimed at researchers and practitioners who encounter limitations with traditional frequentist methods, i.e., methods using $p$-values, especially in cases of null effects and limited sample sizes, the course navigates the advantages of Bayesian approaches over $p$-value based significance testing. The course will highlight the capacity of Bayesian methods to interpret non-significant results meaningfully, which is useful in many areas of perceptual computer graphics. Beyond introducing Bayes theorem, topics include hypothesis evaluation, prior specification, Bayes factors, credible intervals, and how to make the most out of small sample sizes. The course also addresses resources for implementing Bayesian analyses and is accessible to those with a basic understanding of probability and statistics. The course provides a novel contribution to the computer graphics community that can advance experimental methods used to study perception and interaction within graphics.

Additional Key Words and Phrases: Bayesian statistics, Bayes factors, Experimental Analysis

**ACM Reference Format:**
Mirinda Whitaker, Sarah Creem-Regehr, and Bobby Bodenheimer. 2024. Bayesian Statistics: A Practical Introduction for Computer Graphics. 1, 1 (June 2024), 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 OUTLINE

Innovation in interactive 3D computer graphics often relies on user studies to bridge the gap between technical advances and human experience. These user studies can provide quantifiable metrics to guide future design. Yet sometimes interpreting the results of user studies is difficult (e.g., in cases of null effects). This course teaches analytic methods based on Bayes theorem, called Bayesian statistics, that have advantages in many scenarios over more traditional frequentist methods (i.e., using the $p$-value for null hypothesis testing) of analyzing results of studies that are prevalent in computer graphics. For example, Bayesian methods can have utility when working with smaller sample sizes and can provide interpretations of results when traditional null hypothesis significance testing is inadequate to evaluate a research question of interest.

As mentioned, the most common way of analyzing the results of user studies and perceptual experiments in computer graphics is with frequentist statistics, using the $p$-value. Per the American Statistical Association a $p$-value is defined as “the probability under a specified statistical model that a statistical summary of the data (e.g., the sample mean difference between two compared groups) would be equal to or more extreme than its observed value" [Wasserstein and Lazar 2016, p. 131]. In less formal terms, a $p$-value is an index of the inconsistency of our findings with the null

Authors' addresses: Mirinda Whitaker, University of Utah, USA, mirinda.whitaker@utah.edu; Sarah Creem-Regehr, University of Utah, USA, sarah.creem@psych.utah.edu; Bobby Bodenheimer, Vanderbilt University, USA, bobby.bodenheimer@vanderbilt.edu.

hypothesis we have constructed (e.g. no difference between groups, correlation r=0, some percentage representing random chance) in relation to long-run frequencies. $P$-values have been controversial since their inception and their use is still actively debated within the scientific community [Nuzzo 2014]. In practice, $p$-values below some cut-off (often set at the arbitrary $\alpha$ level of .05) are often taken as statistical evidence that an observed difference or association exists and is unlikely to have occurred by chance. However, this interpretation is technically incorrect, and highlights the confusion around the interpretation of $p$-values.

Setting aside the general appropriateness or inappropriateness of $p$-values and null-hypothesis significance testing, in many cases the research questions we want to ask cannot be adequately assessed within a frequentist framework. For example, we may want to determine if people's perception of distance and size is meaningfully different in a real-world room compared to an immersive virtual reality rendering of that room. In many applications, a goal of virtual reality technology is to provide an immersive experience that is as similar to the real-world as possible [Creem-Regehr et al. 2023]. In this scenario, we might hypothesize that with modern computer graphics in immersive displays there would be no meaningful difference in people's perception of distance or size in virtual reality compared to the real-world. Likewise, in evaluating photorealistic virtual humans we are often faced with uncanny valley problems [Carter et al. 2013; Higgins et al. 2021]. Our goal is often to render or animate virtual humans such that they are indistinguishable from real humans. In a frequentist framework we have no way to meaningfully evaluate these hypotheses of no difference. These scenarios are where a Bayesian framework can have immense utility, due to a key difference in how probability is conceptualized.

Bayesian statistics is founded on a different view of probability (subjective probability) and takes a different approach to the analysis of experimental data [van de Schoot et al. 2021]. One approach to understanding Bayesian methods is that they assume prior beliefs (as probabilities) on the hypotheses of an experiment, and then update those beliefs using Bayes theorem with evidence from the data of an experiment, (e.g., a user study, to yield posterior beliefs). Frequentist statistics, on the other hand, are based on long-run frequencies obtained through repeated sampling under near identical conditions to quantify the probability of the data given the null hypothesis. The power of the Bayesian approach is that we can then quantify the odds in favor of one hypothesis versus another, for example the alternative hypothesis in favor of the null hypothesis, and can compute them. These are called Bayes factors [Kass and Raftery 1995]. Another advantage to the Bayesian approach is the idea of credible intervals. A credible interval is a probability statement about the probability (e.g., 90, 95, 99) of an unobserved parameter of a model (e.g., the mean of a population) falling within some interval given the observed data. For example, a 95% credible interval provides a probability statement that given the observed data, there is a 95% probability the unobserved parameter falls within a given credible interval. This may seem like a confidence interval, but a confidence interval is somewhat different, and is often misinterpreted to be like a credible interval. For these reasons, Bayesian methods are becoming are becoming increasingly popular in the cognitive and psychological sciences. However, Bayesian statistical analyses are still uncommon in graphics and graphics-adjacent areas. A reasonable survey found only a few papers employing it [Bodenheimer et al. 2023; Buck et al. 2022, 2020, 2018; Fernandes et al. 2015; Guefrech et al. 2021; Langbehn et al. 2016; Paris et al. 2017; Schmidt et al. 2017; Williams and Peck 2019]. This course aims to educate the community on its strengths and possibilities, and provide additional tools to those designing and analyzing perceptual and user studies.

## 2 RELEVANCE

A goal of user studies in our field is to describe how people perceive and interact with 3D computer graphics. A user study often does this by trying to support a hypothesis regarding how an interface or manipulation will work. It is straightforward to find examples of studies in our literature where no statistically significant results were obtained, i.e., a failure to reject the null hypothesis, e.g., [Gorlewicz et al. 2020; Liu et al. 2020; Sonnenwald et al. 2003]. In a frequentist statistical framework, we can reject the null hypothesis or fail to reject the null hypothesis, but we can never provide *evidence* for the null hypothesis. Yet, oftentimes, especially in user studies, a null or no difference effect can be highly meaningful and is something we want to be able to evaluate and make some kind of claim about. Bayesian statistics offer a complementary perspective on interpreting such data and can provide insights that frequentist methods cannot. For example, with certain assumptions, Bayesian statistics allow one to quantify belief in the null hypothesis and can provide measures of the strength of the evidence supporting the hypothesis of an experiment. Such methods can lead to important insights into perception and interaction in computer graphics that would not otherwise be available.

## 3 TARGET AUDIENCE

The target audience for this course includes researchers and practitioners in computer graphics and related fields who design, run, or analyze user studies. In particular, the course will be useful to those who want to understand modern methods of evaluating the results of user studies. This course is designed to be accessible to anyone with a basic understanding of statistics and probability, such as a CS graduate student.

## 4 TOPICS

**Introduction** In this course we will provide a broad overview of Bayesian statistics and how/when this statistical approach can be used to address commonly encountered issues (e.g. null effects, limited sample sizes) that arise for those analyzing data from user studies on computer graphics.

**Bayesian basics** Here we will introduce Bayes theorem, which is the basis of Bayesian statistics. Bayes theorem provides a different view of probability (as opposed to frequentist probability), which allows us to evaluate a hypothesis given our data $P(H_0|data)$, instead of being limited to assessing only the probability of the data given the null hypothesis $P(data|H_0)$.

**Bayesian data analysis** Bayes theorem is a powerful theorem with many applications in neat and discrete applications like those that will be detailed in the prior section. However, this theorem can also be harnessed alongside recent computational advancements and Monte Carlo methods to tackle messier data and various statistical analyses commonly used in the behavioral sciences. Bayesian statistics is a powerful statistical framework that draws on a particular view of probability and inference. Frequentist statistics relies on the assumption of repeated sampling under near-identical conditions to obtain long-run frequencies and infer *point estimates* about parameters, whereas Bayesian statistics incorporates a subjective belief quantification that gets updated as data are observed to infer about the *distributions* of parameters. In Bayesian statistics, we quantify our beliefs and our certainty/uncertainty in those beliefs through priors. While incorporating subjective beliefs into statistical analyses can feel odd, we get to decide the amount of information we encode into a prior. In fact in many cases, we can specify a completely uninformative prior (e.g. a uniform distribution from $-\infty$ to $+\infty$), and close to identically replicate frequentist analyses. In this section, we will discuss the role of priors and some general guidelines about how to specify priors. We will also discuss Bayes factors and credible intervals.

**When/why to use Bayesian Statistics?** We will provide an overview of common data scenarios where Bayesian data analysis has advantages. One such example is a common goal of improving the *perceptual fidelity* of computer graphics, or the likelihood that people will perceive and act with computer graphics as they would with real-world objects or environments [Pointon et al. 2018; Stefanucci et al. 2015]. Differences in perceptual fidelity have been a long-time topic of study for space perception in virtual and augmented reality but modern technologies and measures have reduced some of these effects [Buck et al. 2018; Creem-Regehr et al. 2023]. This goal of equivalence in perceptual fidelity may ultimately lead to a prediction of a null result, i.e., no difference between a real and virtual environment condition. In this circumstance, Bayesian statistics would be a preferable method to evaluate the strength of the evidence supporting the null hypothesis. Likewise, other conditions that may be predicted to influence perception of graphics (e.g., rendered shadows, object characteristics) may not have the expected effects and the strength of the effects could be more effectively evaluated with Bayesian methods [Buck et al. 2020]. In the following two sections, we present two common data scenarios (null effects and limited sample size) to highlight the utility of a Bayesian approach.

**Hypothesis evaluation** Here we will highlight that a Bayesian approach gives one tools to evaluate the null hypothesis and say something about results that would otherwise be just $p > .05$. We will also show how Bayes factors can be used to evaluate the likelihood of a given hypothesis and how credible intervals can be used to index the range of possible effect sizes.

**Limited sample sizes** In this section, we will provide an overview of how building in prior information can be useful with small samples [Lee and Song 2004]. More specifically, we will highlight how with appropriate and informed priors, a Bayesian approach can allow one to make the most out of small sample sizes that may arise out of pilot/feasibility testing or hard to recruit populations.

**Resources** Finally, we will provide some recommendations for software that can be used to implement Bayesian analyses (e.g., `brms` and STAN for R users; PyMC for Python users), as well as resources for those who want to learn more about Bayesian analyses (e.g., *Doing Bayesian Data Analysis* [Kruschke 2014]).

## 5 SYLLABUS

We propose a short (1.5 hour) course according to the following schedule:

(1) Introduction - Sarah Creem-Regehr (5 minutes)
(2) Bayesian basics - Bobby Bodenheimer (15 minutes)
(3) Bayesian data analysis - Mirinda Whitaker (15 minutes)
(4) When/why to use Bayesian Statistics? - Sarah Creem-Regehr (10 minutes)
(5) Hypothesis evaluation - Mirinda Whitaker (10 minutes)
(6) Limited sample sizes - Mirinda Whitaker (10 minutes)
(7) Resources - Mirinda Whitaker (5 minutes)
(8) Conclusions/Summary - Bobby Bodenheimer (5 minutes)
(9) Q&A - All (15 minutes)

## 6 RELATED COURSES

To our knowledge, no similar course has been offered at SIGGRAPH. Bayesian decision theory has been covered in prior Siggraph courses, e.g., the 2004 Siggraph course by Aaron Hertzmann [Hertzmann 2004] or the discussion of it in the more recent 2018 Siggraph course “Applications of Vision Science to Virtual and Augmented Reality” by Patney and colleagues [Patney et al. 2018]. Bayesian decision theory uses Bayes theorem to determine what an optimal decision would be based on the available evidence and is distinct from the use of Bayesian statistics as we have described it here. Likewise, courses that have dealt with user studies such as the 2008 course by Ferwerda on psychophysics [Ferwerda

2008] or the 2009 course by Sundstedt and colleagues [Sundstedt et al. 2009] have not focused on the analysis of experimental data.

## 7 AUTHOR BIOGRAPHIES

**Mirinda Whitaker** Department of Psychology, University of Utah, Salt Lake City, UT, USA.

**Email:** mirinda.whitaker@utah.edu

Mirinda Whitaker is a Psychology PhD student at the University of Utah, jointly in Cognition and Neural Science and Quantitative Psychology. She received her BS in 2019 and her MS in 2022 both in Psychology. Her methodological expertise is in Bayesian statistics and her program of research aims to develop novel statistical/quantitative methods that can be applied to further understand how we perceive and interact with the world around us. She has 6 peer-reviewed publications (3 first-author) across cognition and perception journals (e.g., *Attention, Perception, & Psychophysics*). She has been invited to lend her expertise as an ad-hoc reviewer for 11 manuscripts at numerous psychology journals (e.g., *Psychonomic Bulletin and Review*) and computer science conferences (e.g., *IEEE VR*). She has extensive statistical teaching experience, serving as a teaching assistant for 1 undergraduate, 2 honors undergraduate, 2 introductory graduate, and 2 upper-level graduate statistical courses. She has served as a statistical consultant on several grants and has given tutorial talks on topics such as Bayesian statistics to faculty and graduate students at the University of Utah. She has presented her work using Bayesian statistics to analyze experimental data (at venues such as the *Society for Affective Science Conference* and the *Psychonomic Society Conference*). She recently presented a novel Bayesian model at a top quantitative psychology conference (*Modern Modeling Methods: M3*) and has a manuscript reporting on this model under-review at *Psychological Methods*.

**Sarah Creem-Regehr** Department of Psychology, University of Utah, Salt Lake City, UT, USA.

**Email:** sarah.creem@psych.utah.edu

Sarah Creem-Regehr is a Professor in the Psychology Department at the University of Utah, where she has been since 2000. She also holds faculty appointments in the School of Computing and the Neuroscience program at the University of Utah. Her research examines how humans perceive, learn, and navigate spaces in natural, virtual, and visually impoverished environments. Her work in computer graphics and virtual environments has contributed to solutions to improve the utility of virtual environment applications by studying human perception and performance. She co-authored the book *Visual Perception from a Computer Graphics Perspective*. She served as a conference program chair for IEEE VR in 2022, a program chair for the ACM Symposium on Applied Perception in 2008, Associate Editor for *Psychonomic Bulletin & Review*, *Journal of Experimental Psychology: HPP*, and *Quarterly Journal of Experimental Psychology*, and is currently Editor-in-Chief of *Cognitive Research: Principles and Implications*.

**Bobby Bodenheimer** Department of Computer Science, Vanderbilt University, Nashville, TN, USA.

**Email:** bobby.bodenheimer@vanderbilt.edu

Bobby Bodenheimer is a professor of Computer Science at Vanderbilt University, where he has been since 2000. His research currently is in the area of virtual and augmented reality, where he focuses on how people perceive and learn in virtual and augmented spaces. His work has used Bayesian statistics to evaluate the performance of methods in virtual reality when evaluated by user studies. He is the current Editor-in-Chief of the ACM *Transactions on Applied Perception*. He was one of the technical program chairs of IEEE VR in 2023, a conference program chair for IEEE VR in 2022, a program chair for the ACM Symposium on Applied Perception in 2009, a conference chair for the ACM Symposium on Computer Animation in 2009, and a conference chair for the Symposium on Applied Perception in Graphics and Visualization in 2008, and previously served as an Associate Editor for the IEEE *Transactions on Visualization and Computer Graphics*.

## REFERENCES


Bobby Bodenheimer, Haley Adams, Mirinda Whitaker, Jeanine Stefanucci, and Sarah Creem-Regehr. 2023. Perceiving Absolute Distance in Augmented Reality Displays with Realistic and Non-realistic Shadows. In *ACM Symposium on Applied Perception 2023* (Los Angeles, CA, USA) *(SAP '23)*. Association for Computing Machinery, New York, NY, USA, Article 2, 9 pages. https://doi.org/10.1145/3605495.3605800

Lauren E. Buck, Soumyajit Chakraborty, and Bobby Bodenheimer. 2022. The Impact of Embodiment and Avatar Sizing on Personal Space in Immersive Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics* 28, 5 (2022), 2102–2113. https://doi.org/10.1109/TVCG.2022.3150483

Lauren E. Buck, Sohee Park, and Bobby Bodenheimer. 2020. Determining Peripersonal Space Boundaries and Their Plasticity in Relation to Object and Agent Characteristics in an Immersive Virtual Environment. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 332–342. https://doi.org/10.1109/VR46266.2020.00053

Lauren E. Buck, Mary K. Young, and Bobby Bodenheimer. 2018. A Comparison of Distance Estimation in HMD-Based Virtual Environments with Different HMD-Based Conditions. *ACM Transactions on Applied Perception (TAP)* 15, 3, Article 21 (July 2018), 15 pages. https://doi.org/10.1145/3196885

Elizabeth J. Carter, Moshe Mahler, and Jessica K. Hodgins. 2013. Unpleasantness of animated characters corresponds to increased viewer attention to faces. In *Proceedings of the ACM Symposium on Applied Perception* (Dublin, Ireland) *(SAP '13)*. Association for Computing Machinery, New York, NY, USA, 35–40. https://doi.org/10.1145/2492494.2502059

Sarah H Creem-Regehr, Jeanine K Stefanucci, and Bobby Bodenheimer. 2023. Perceiving distance in virtual reality: theoretical insights from contemporary technologies. *Philosophical Transactions of the Royal Society B* 378, 1869 (2023), 20210456.

Ajoy S. Fernandes, Ranxiao Frances Wang, and Daniel J. Simons. 2015. Remembering the physical as virtual: source confusion and physical interaction in augmented reality. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception* (Tübingen, Germany) *(SAP '15)*. Association for Computing Machinery, New York, NY, USA, 127–130. https://doi.org/10.1145/2804408.2804423

James A. Ferwerda. 2008. Psychophysics 101: how to run perception experiments in computer graphics. In *ACM SIGGRAPH 2008 Classes* (Los Angeles, California) *(SIGGRAPH '08)*. Association for Computing Machinery, New York, NY, USA, Article 87, 60 pages. https://doi.org/10.1145/1401132.1401243

Jenna L. Gorlewicz, Jennifer L. Tennison, P. Merlin Uesbeck, Margaret E. Richard, Hari P. Palani, Andreas Stefik, Derrick W. Smith, and Nicholas A. Giudice. 2020. Design Guidelines and Recommendations for Multimodal, Touchscreen-based Graphics. *ACM Trans. Access. Comput.* 13, 3, Article 10 (aug 2020), 30 pages. https://doi.org/10.1145/3403933

Fatma Ben Guefrech, Florent Berthaut, Patricia Plénacoste, Yvan Peter, and Laurent Grisoni. 2021. Revealable Volume Displays: 3D Exploration of Mixed-Reality Public Exhibitions. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 31–39. https://doi.org/10.1109/VR50410.2021.00023

Aaron Hertzmann. 2004. Introduction to Bayesian learning. In *ACM SIGGRAPH 2004 Course Notes* (Los Angeles, CA) *(SIGGRAPH '04)*. Association for Computing Machinery, New York, NY, USA, 22–es. https://doi.org/10.1145/1103900.1103922

Darragh Higgins, Donal Egan, Rebecca Fribourg, Benjamin Cowan, and Rachel McDonnell. 2021. Ascending from the valley: Can state-of-the-art photorealism avoid the uncanny?. In *ACM Symposium on Applied Perception 2021* (Virtual Event, France) *(SAP '21)*. Association for Computing Machinery, New York, NY, USA, Article 7, 5 pages. https://doi.org/10.1145/3474451.3476242

Robert E. Kass and Adrian E. Raftery. 1995. Bayes Factors. *J. Amer. Statist. Assoc.* 90, 430 (1995), 773–795. https://doi.org/10.1080/01621459.1995.10476572 arXiv:http://amstat.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476572

John Kruschke. 2014. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan.* Academic Press.

Eike Langbehn, Tino Raupp, Gerd Bruder, Frank Steinicke, Benjamin Bolte, and Markus Lappe. 2016. Visual blur in immersive virtual environments: does depth of field or motion blur affect distance and speed estimation?. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (Munich, Germany) *(VRST '16)*. Association for Computing Machinery, New York, NY, USA, 241–250. https://doi.org/10.1145/2993369.2993379

Sik-Yum Lee and Xin-Yuan Song. 2004. Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate behavioral research* 39, 4 (2004), 653–686.

Shi-Hong Liu, Pai-Chien Yen, Yi-Hsuan Mao, Yu-Hsin Lin, Erick Chandra, and Mike Y. Chen. 2020. HeadBlaster: a wearable approach to simulating motion perception using head-mounted air propulsion jets. *ACM Trans. Graph.* 39, 4, Article 84 (aug 2020), 12 pages. https://doi.org/10.1145/3386569.3392482

Rigina Nuzzo. 2014. Statistical errors: P values, the'gold standard'of statistical validity, are not as reliable as many scientists assume. *Nature* 506, 7487 (2014), 150–153.

Richard Paris, Miti V. Joshi, Qiliang He, Gayathri Narasimham, Timothy P. McNamara, and Bobby Bodenheimer. 2017. Acquisition of Survey Knowledge using Walking in Place and Resetting Methods in Immersive Virtual Environments. In *Proceedings of the ACM Symposium on Applied Perception* (Cottbus, Germany) *(SAP '17)*. ACM, New York, NY, USA.

Anjul Patney, Marina Zannoli, Joohwan Kim, Robert Konrad, Frank Steinicke, and Martin S Banks. 2018. Applications of vision science to virtual and augmented reality. In *ACM SIGGRAPH 2018 Courses*. 1–28.

Grant Pointon, Chelsey Thompson, Sarah Creem-Regehr, Jeanine Stefanucci, Miti Joshi, Richard Paris, and Bobby Bodenheimer. 2018. Judging action capabilities in augmented reality. In *Proceedings of the 15th ACM Symposium on Applied Perception*. 1–8.

Susanne Schmidt, Gerd Bruder, and Frank Steinicke. 2017. Moving Towards Consistent Depth Perception in Stereoscopic Projection-based Augmented Reality. In *ICAT-EGVE 2017 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, Robert W. Lindeman, Gerd Bruder, and Daisuke Iwai (Eds.). The Eurographics Association. https://doi.org/10.2312/egve.20171354

Diane H. Sonnenwald, Mary C. Whitton, and Kelly L. Maglaughlin. 2003. Evaluating a scientific collaboratory: Results of a controlled experiment. *ACM Trans. Comput.-Hum. Interact.* 10, 2 (jun 2003), 150–176. https://doi.org/10.1145/772047.772051

Jeanine K Stefanucci, Sarah H Creem-Regehr, William B Thompson, David A Lessard, and Michael N Geuss. 2015. Evaluating the accuracy of size perception on screen-based displays: Displayed objects appear smaller than real objects. *Journal of Experimental Psychology: Applied* 21, 3 (2015), 215–223.

Veronica Sundstedt, Mary Whitton, and Marina Bloj. 2009. The whys, how tos, and pitfalls of user studies. In *ACM SIGGRAPH 2009 Courses* (New Orleans, Louisiana) *(SIGGRAPH '09)*. Association for Computing Machinery, New York, NY, USA, Article 25, 205 pages. https://doi.org/10.1145/1667239.1667264

Rens van de Schoot, Sarah Depaoli, Ruth King, Bianca Kramer, Kaspar Märtens, Mahlet G Tadesse, Marina Vannucci, Andrew Gelman, Duco Veen, Joukje Willemsen, et al. 2021. Bayesian statistics and modelling. *Nature Reviews Methods Primers* 1, 1 (2021), 1.

Ronald L Wasserstein and Nicole A Lazar. 2016. The ASA statement on p-values: context, process, and purpose. , 129–133 pages.

Niall L. Williams and Tabitha C. Peck. 2019. Estimation of Rotation Gain Thresholds Considering FOV, Gender, and Distractors. *IEEE Transactions on Visualization and Computer Graphics* 25, 11 (2019), 3158–3168. https://doi.org/10.1109/TVCG.2019.2932213