TU Dortmund

Introductory Case Studies

Project 2: Comparison of k Distributions Using the Example of an Anchoring Effect Experiment

Lecturers:
Prof. Dr. Philipp Doebler
Dr. Dennis Dobler, JProf
Dr. Susanne Frick
Loreen Sabel

Author: Md Omar Faruk

Group number: 01

Group members: Tazin Rahman

June 1, 2025

Contents

1 Introduction

The anchoring effect is a cognitive bias when people base their numerical judgments unduly on an initial number, known as the anchor. In areas like pricing, negotiations, and decision-making, where the presence of an anchor can greatly affect final estimates, this phenomenon has been widely recognized.

According to earlier studies, judgment can be impacted by even minute aspects of the anchor, such its accuracy. More decimal places in an anchor, for instance, have been demonstrated to produce estimates that are closer to the anchor than rounded numbers. The psychological significance that people attach to particular numerical clues is demonstrated by these findings.

This project examines information from a replication study of the anchoring effect in response to growing concerns regarding replicability in psychological science. We investigate the differences in participant estimations as a function of anchor precision and motivational framing using Chandler (2015)'s responses. Nine distinct estimating tasks are included in the data, which offer a chance to evaluate the generalizability of earlier findings in other settings.

The aim of this analysis is to determine whether this independent dataset exhibits the previously documented impacts of anchor precision on judgmental estimates.

Section 2 presents the problem statement, outlining the core research question and objectives related to the precision effect in numerical anchoring. Section 3 describes the statistical methods used to analyze the data, including both parametric and non-parametric approaches. Section 4 provides a detailed statistical analysis, including assumption checks, hypothesis testing, and graphical summaries. Finally, Section 5 summarizes the key findings and discusses their implications in the context of the original study by Janiszewski and Uy (2008).

2 Problem statement

This project uses data from an experimental investigation that was supposed to examine how judgment is affected by numerical anchoring. Based on Janiszewski and Uy's (2008) research, the study focuses on the precision effect, which suggests that estimates may be more biased by precise numerical anchors than by rounded ones. A 2×2 factorial design was used to randomly assign participants to one of four groups based on two independent variables: magnitude (low vs. high numerical anchor) and anchortype (precise vs. round). At first, information was gathered from 126 individuals. The condition of independent judgment was broken by six participants, who were subsequently removed from the study for discussing their responses with one another. Therefore, there are 30 people in each experimental group, and the final dataset has 120 observations.

2.1 Description of the project

Nine common consumer goods—pen, protein drink, lebron, slidy, cheese, figurine, TV, beach home, and number—were given to the participants to estimate their values. All of these continuous item-specific variables are measured using a ratio scale. The average relative deviation between participants' estimates and the anchors they were shown is captured by the main dependent variable, MeanRelativeDifference. In the analysis, anchoring influence is primarily measured by this variable.

Several demographic and metadata variables are included:

- Drop: A binary indicator (typically 0 or 1) signaling whether a row should be excluded from analysis. Rows marked with 1 in the "drop" column were removed during preprocessing to ensure data quality and adherence to experimental design.
- ParticipantID: A unique identifier for each participant.
- Age: A numerical variable representing the participant's age.
- Gender: A categorical variable indicating male or female.
- Year: The year in which the participant took part in the study.

2.2 Objective of the project

This experiment had two objectives: to statistically evaluate the impact of Anchortype, magnitude, and their interplay on numerical estimations, and to replicate important discoveries on the precision effect. Through controlled replication, this initiative assesses the validity of previous psychological findings and advances our understanding of cognitive biases in judgment.

3 Statistical methods

3.1 Statistical Measures

Hypothesis Testing

The process of determining whether to accept or reject a hypothesis regarding an unknown parameter within the distribution of a random variable is known as a statistical test. Two categories of hypotheses exist: the null hypothesis, denoted as H_0, is the main hypothesis, whereas H_A stands for the alternative hypothesis. H_A is rejected if H_0 is accepted, while H_0 is rejected if H_A is rejected (Rasch et al., 2019, page 39).

There are two kinds of errors that can happen in statistical testing. When the null hypothesis is true but is mistakenly rejected, this is known as a Type I error. The significance level, denoted by α, represents the likelihood of making a Type I error and is usually adjusted to 0.05 or 0.01 before the test. More solid evidence is required to reject the null hypothesis, as indicated by a smaller significance level. When the null hypothesis is incorrect but not rejected, this is known as a type II error, and it is represented by the symbol β. Since a rise in α typically results in a decrease in β, and vice versa, type I and type II mistakes cannot occur in the same test. One way to reduce the possibility of both kinds of errors is to increase the sample size.

Comparing the p-value with the significance threshold α is a popular method for drawing conclusions in hypothesis testing. The likelihood of finding a result that is even less likely under the null hypothesis is represented by the p-value, commonly referred to as the observed significance level. The null hypothesis is rejected if the p-value is less than the selected significance threshold α (Black, 2009, pages 298-302).

Two-way ANOVA Test

Understanding the effects of several independent variables (factors) on a dependent variable is frequently required in experimental and psychological research. A statistical technique created specifically for this purpose is a two-way ANOVA, sometimes referred to as a factorial ANOVA. The major impacts of each factor are evaluated by two-way ANOVA, which also looks for interaction effects, or whether the influence of one component is dependent on the level of another. This is in contrast to one-way ANOVA, which only examines one factor at a time.

The hypotheses for the factorial ANOVA test are defined as follows. For each factor and its interaction, the **null hypothesis** (H_0) assumes that all group means are equal.

For Factor A (with R levels):

$$H_0^A : \mu_{1\cdot} = \mu_{2\cdot} = \cdots = \mu_{R\cdot}$$

For Factor B (with C levels):

$$H_0^B : \mu_{\cdot 1} = \mu_{\cdot 2} = \cdots = \mu_{\cdot C}$$

For the Interaction between Factor A and Factor B:

$$H_0^{AB} : \text{There is no interaction effect between the two factors.}$$

The **alternative hypothesis** (H_A) suggests that at least one group mean differs or that an interaction exists, indicating that the combined effects of the factors are not purely additive.

In a two-way ANOVA, an F-statistic is calculated for each main effect and the interaction. The F-statistic compares the mean square of a factor (or interaction) to the residual (error) mean square.

1. Mean Squares (MS) For each component, the mean square is calculated by dividing the sum of squares (SS) by its corresponding degrees of freedom (df):

$$\begin{aligned}
MS_A &= \frac{SS_A}{df_A} = \frac{SS_A}{R-1} \\
MS_B &= \frac{SS_B}{df_B} = \frac{SS_B}{C-1} \\
MS_{AB} &= \frac{SS_{AB}}{df_{AB}} = \frac{SS_{AB}}{(R-1)(C-1)} \\
MS_R &= \frac{SS_R}{df_R} = \frac{SS_R}{R \cdot C \cdot (N-1)}
\end{aligned}$$

2. F-Statistics The F-statistics for each effect are then computed as:

$$F_A = \frac{MS_A}{MS_R}$$
$$F_B = \frac{MS_B}{MS_R}$$
$$F_{AB} = \frac{MS_{AB}}{MS_R}$$

Each F-statistic follows an F-distribution under the null hypothesis, with numerator and denominator degrees of freedom corresponding to the respective MS terms.

The observed F-value for each factor or interaction is compared to a critical value from the F-distribution, which is determined by the degrees of freedom for the numerator and denominator (e.g., df_A, df_R) and the selected significance level α. If the observed F-value is greater than the critical value, the null hypothesis is rejected, suggesting a statistically significant effect. Otherwise, the null hypothesis is retained, indicating that the group means do not differ significantly. (Field, 2012, page 180-181)

Shapiro-Wilk Test

The Shapiro-Wilk test is a statistical test used to assess whether a sample comes from a normally distributed population.

- **Null Hypothesis (H_0):** The data is normally distributed.
- **Alternative Hypothesis (H_A):** The data is not normally distributed.

The Shapiro-Wilk test statistic W is defined as:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Where:

- $x_{(i)}$: The i-th order statistic (i.e., the i-th smallest value in the sample)
- $\bar{x}$: The sample mean
- a_i: Constants computed from the expected values of order statistics of a normally distributed sample and the covariance matrix

- n: The sample size
- W: The Shapiro-Wilk test statistic, ranging between 0 and 1
- If W is close to 1, the sample is likely to be normally distributed.
- A small W indicates deviation from normality.
- The corresponding p-value is used to determine significance:
  - If $p < 0.05$, reject the null hypothesis: the data is not normally distributed.
  - If $p \geq 0.05$, fail to reject the null hypothesis: the data is normally distributed. (Moore et al. 2012, p. 125)

Levene's Test Summary

Levene's test is used to assess the equality of variances across two or more groups (i.e., homogeneity of variances).

Hypotheses:

- Null Hypothesis (H_0): The variances are equal across groups.
- Alternative Hypothesis (H_A): At least one group has a different variance.

Procedure:

1. Compute the absolute deviations from the group mean (or median):

$$z_{ij} = |x_{ij} - \bar{x}_i|$$

2. Conduct a one-way ANOVA on the z_{ij} values.
3. Evaluate the p-value:
   - If $p < 0.05$, reject H_0: variances are not equal.
   - If $p \geq 0.05$, fail to reject H_0: variances are equal (Montgomery et al., 2012).

3.2 Statistical Plots

Histogram

One kind of graph that shows the distribution of numerical data is a histogram. The data is grouped into continuous intervals, sometimes known as bins, and a bar is drawn for each interval. Numerical ranges are represented by the x-axis, and the frequency of data inside each range is shown by the y-axis. Each bar's height in a histogram is determined by the formula $h_j = \frac{f_j}{d_j}$, where dj is the bar's width and fj is the range's frequency (Schomaker, 2016, Page 227).

A continuous variable's histogram is said to be symmetric if its mean and median are equal. A histogram that is right-skewed shows that the mean is pulled to the right by a small number of extreme values, while numerous small values are positioned to the left. A left-skewed histogram, on the other hand, indicates that the mean value is pulled to the left by a few outliers and that there are numerous minor data points to the right.

Box Plots

An overview of a continuous variable is given by a box plot. The minimum, first quartile, median, third quartile, and maximum values are among the important summary statistics that are shown. The first quartile is represented by the box's lower hinge, the median value is shown by the horizontal line inside the box, and the third quartile is represented by the upper hinge. Outliers are defined as values that deviate from the norm.

4 Statistical analysis

This section describes the statistical techniques used to examine how participants' numerical estimates—more especially, their mean relative difference from the provided anchor—were impacted by anchor precision (Anchortype) and incentive level (magnitude). All analyses were carried out in a Jupyter Notebook environment using Python (version 3.11.4) and the pandas (version 1.5.3) and sklearn (version 1.3.0) libraries. For all inferential tests, the significance threshold (α) was set at 0.05.

4.1 Descriptive Statistics and Data Exploration:

The Mean Relative Difference, which was determined for each participant by averaging (Participant's Estimate - Actual Anchor Value) / Actual Anchor Value across nine distinct estimation items, served as the dependent variable. Estimates below the anchor (undershoot) are indicated by negative values on this measure, estimates above the anchor (overshoot) are indicated by positive values, and estimates closer to the stated anchor are represented by values near zero.

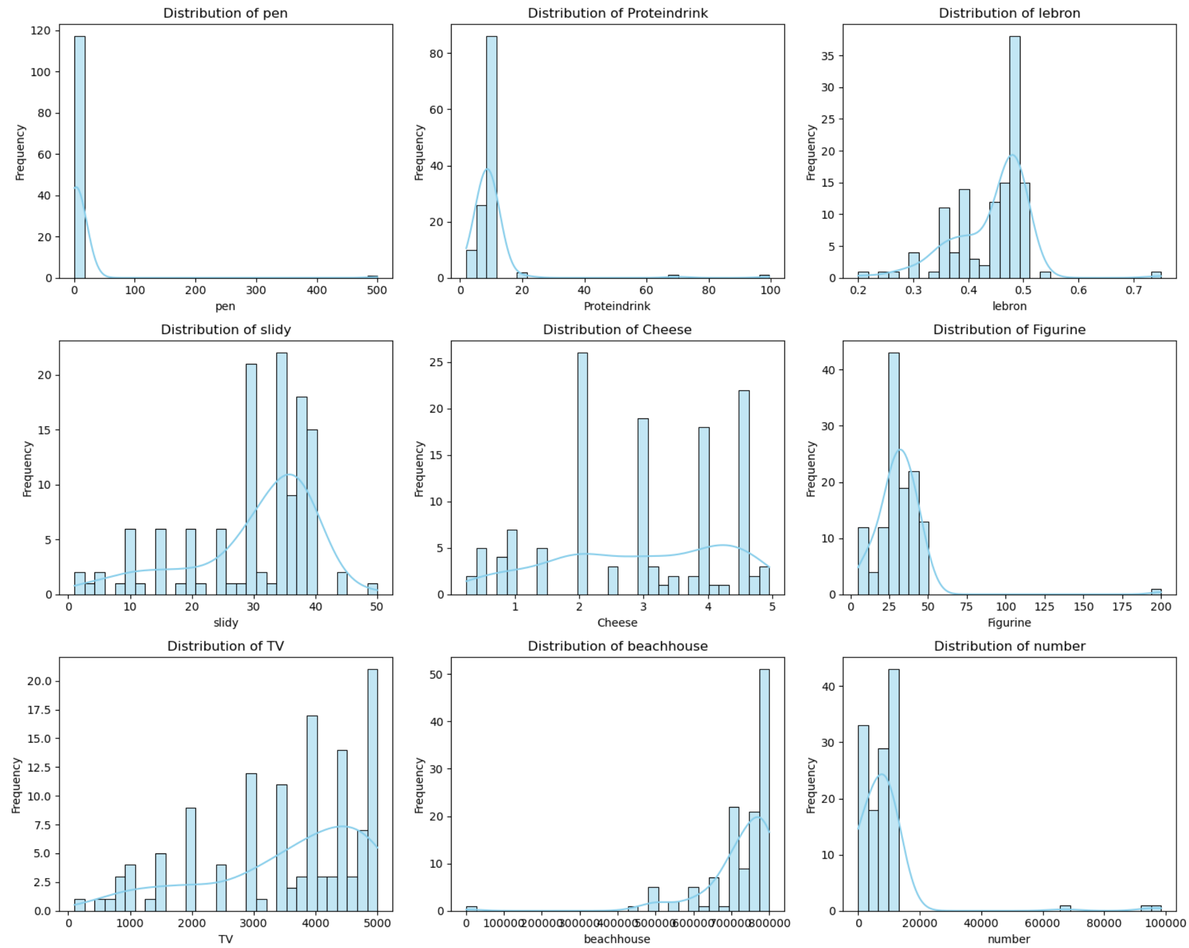

Figure 1: Boxplots of Raw Estimates for Nine Items

Boxplots were used to perform a first investigation of the raw estimates for each of the nine elements (see Figure 1). The x-axis represents the range of estimates given by participants for that specific item, and the y-axis represents the frequency of that range.

4.2 Detailed Presentation of ANOVA Results:

A 2×2 (*Anchor Type* $\times$ *Motivation Level*) between-subjects factorial analysis of variance (ANOVA) was conducted to examine the effects of anchor precision and motivation on mean relative difference from the anchor.

Source	sum_sq		df	F	PR(>F)	
C(Anchortype)	7.76	$\times 10^{-2}$	1.00	8.55	4.16	$\times 10^{-3}$
C(magnitude)	1.43		1.00	1.57	2.42	$\times 10^{-23}$
C(Anchortype):C(magnitude)	1.13	$\times 10^{-4}$	1.00	1.25	9.11	$\times 10^{-1}$
Residual	1.05		1.16			

Table 1: ANOVA Table

Results revealed a significant main effect of motivation level, $F(1, 116) = 157.06$, $p < .05$. There was also a significant main effect of anchor type, $F(1, 116) = 8.55$, $p < .05$.

However, the interaction between anchor type and magnitude was not significant, $F(1, 116) = 0.01$, $p = .911$, indicating that the effect of magnitude on the outcome was consistent regardless of the anchor type used.

4.3 Precision Level vs. Anchoring Effect

The box plot titled "Effect of Anchortype on Mean Relative Difference" visually summarizes the relationship between the type of anchor (round vs. precise) and the mean relative difference of estimates from the anchor.

- X-axis ("Anchortype"): Shows the two levels of anchor precision: "round" and "precise."
- A Y-axis ("Mean Relative Difference"): Represents how much the participants' estimates differed from the given anchor, relative to the anchor value. Since participants were asked to estimate below the anchor, these values are negative.

The box plot visually suggests that using a precise anchor leads to estimates that are closer to the anchor (less negative relative difference) compared to using a round anchor. This provides visual support for the "precision effect."

4.4 Checks of Underlying Assumptions:

- **Independence of Observations:** Observations were assumed to be independent as participants were [e.g., randomly assigned to one of the four experimental conditions and data were collected individually].

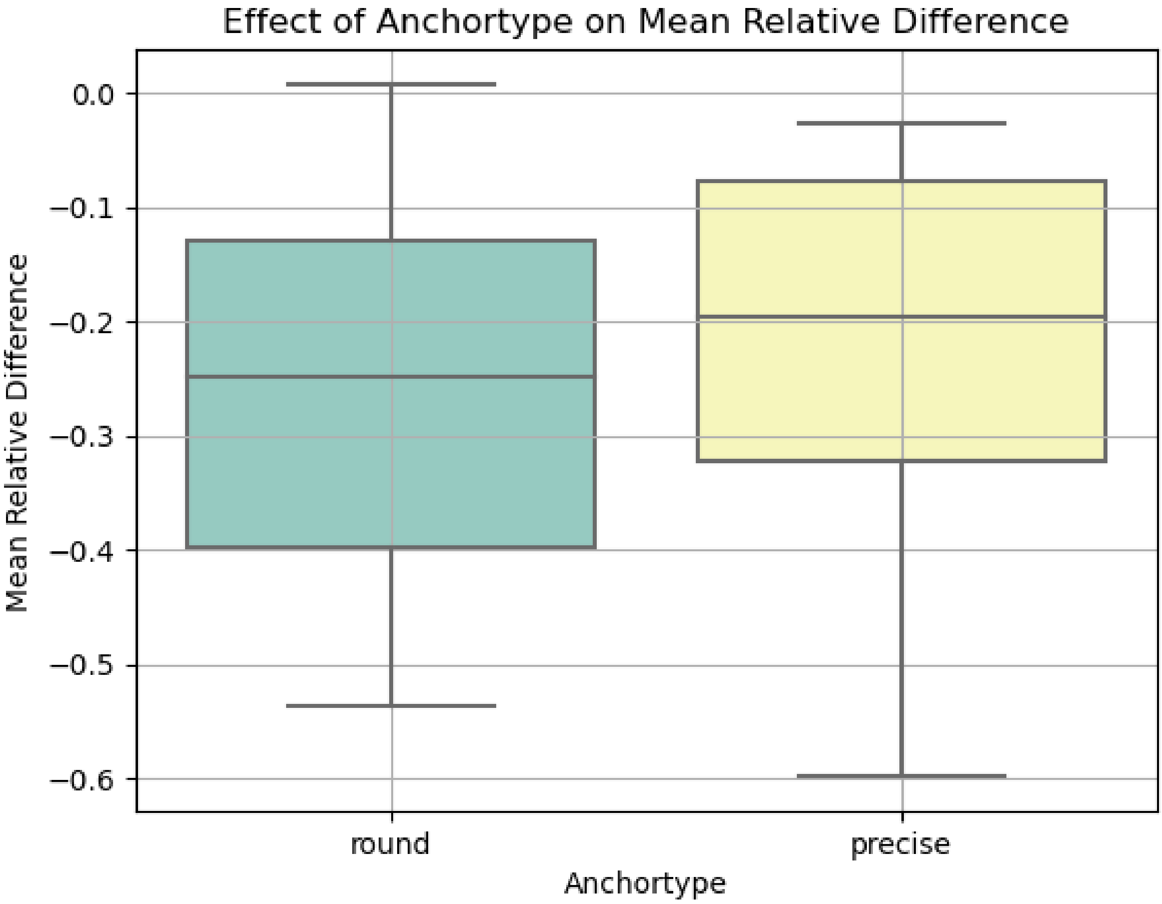

Figure 2: Effect of Anchortype on Mean Relative Difference

- **Homogeneity of Variances:** Levene's test was used to examine the hypothesis of homogeneity of variances among the four experimental groups. With F(3, 116) = 2.113, p =.102, Levene's test was not significant, indicating that the variances were identical across groups and that the homogeneity of variance assumption was satisfied.
- **Normality of Residuals:** The normality of the residuals from the ANOVA model was evaluated by visual inspection of a histogram. The Shapiro-Wilk test indicated that the residuals significantly deviated from a normal distribution, W = 0.95, p = .0003. While ANOVA is considered relatively robust to violations of normality, particularly with balanced designs (as is the case here with n=30 per cell) and sufficient sample size (N=120), this result suggests that the p-values from the ANOVA should be interpreted with some caution.

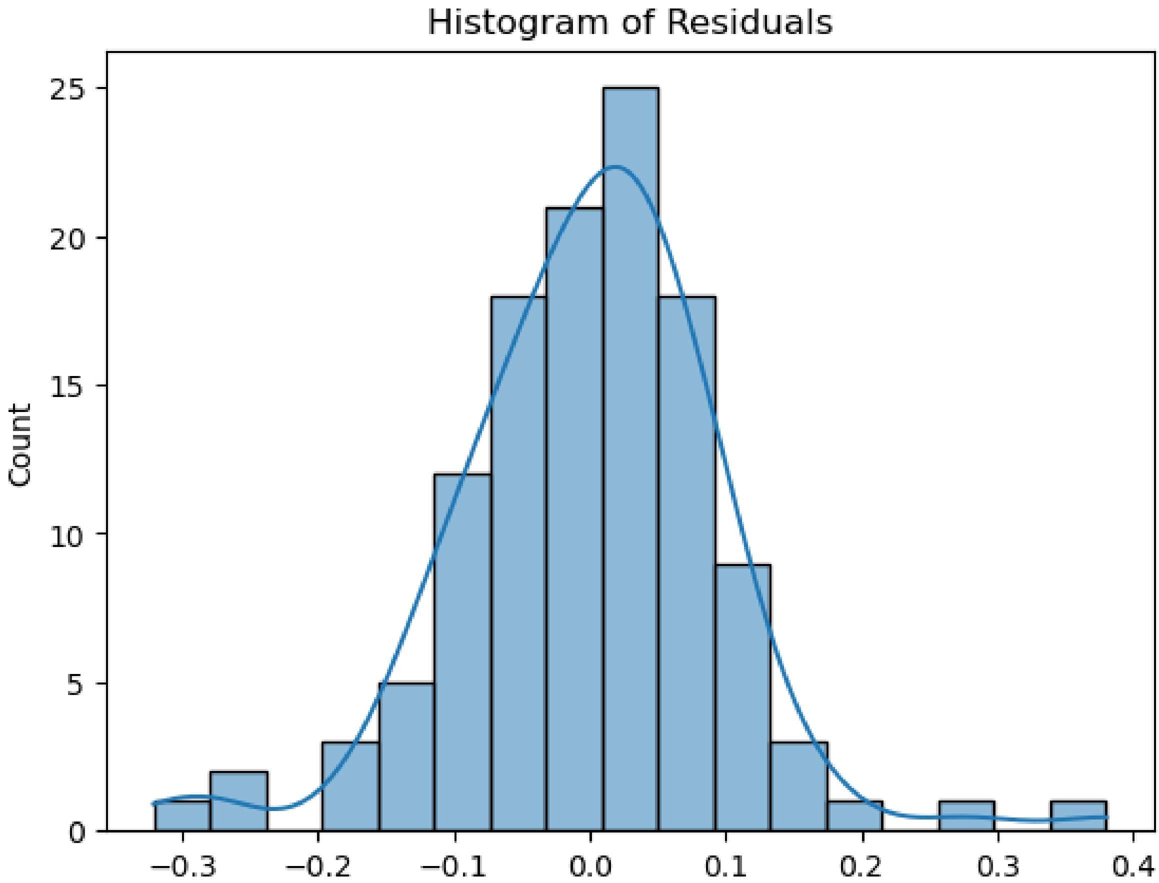

Figure 3: Histogram of Residuals

5 Summary

This project aimed to investigate the influence of anchor precision and motivation to underestimate on numerical estimation behavior, specifically examining how these factors affect the degree of adjustment from a given anchor. A smaller (less negative, closer to zero) mean relative difference indicates estimates closer to the anchor, while a larger (more negative) difference indicates estimates further away.

The analysis of variance (ANOVA) revealed several key findings:

1. Main Effect of Anchortype: A statistically significant main effect of Anchortype was found. This indicates that the precision of the numerical anchor significantly influences how much participants adjust their estimates. As visually depicted in the provided box plot, participants exposed to precise anchors (e.g., 9.8g, 39.75) tended to make estimates that were, on average, closer to the anchor (less negative mean relative difference) compared to those exposed to round anchors (e.g., 10g, 40).

2. Main Effect of Magnitude: A highly statistically significant main effect of Magnitude (motivation to underestimate) was also observed $p = 2.81 \times 10^{-23}$. This demonstrates that the level of motivation (stronger vs. slight underestimation) substantially impacts how far estimates adjust from the anchor.

3. Interaction Effect: Crucially, the interaction effect between Anchortype and Magnitude was not statistically significant (p = 0.956). This implies that the effect of anchor precision on numerical estimates does not significantly depend on the level of motivation, and vice versa. Both factors appear to influence judgments independently.

Comparison with Janiszewski and Uy (2008): Janiszewski and Uy (2008) concluded that "higher precision led to estimates closer to the anchor." Our findings are in direct agreement with this core conclusion. The significant main effect of Anchortype, showing estimates closer to precise anchors, mirrors their original discovery.

Can we say it is a successful replication? Yes, we can consider this a successful replication regarding the primary hypothesis concerning the precision effect. The consistency of the results from Chandler's data with Janiszewski and Uy's (2008) key finding strengthens the reliability and generalizability of the precision anchoring effect. In the context of the replication crisis, successfully replicating a finding with independent data is a valuable contribution to the psychological literature.

Practical Applications and Relevance of the Anchor Effect: The anchoring effect, and specifically the precision effect, has highly significant practical applications and is undoubtedly practically relevant. As detailed in the project background:

- Negotiation Situations: In bargaining, an initial offer with a precise number (e.g., "99,998" instead of "100,000") can anchor the other party's counter-offers closer to that precise value, potentially leading to a more favorable final agreement.
- General Decision-Making: In any scenario where people must estimate a numerical value under uncertainty (e.g., project timelines, resource allocation, market forecasts), the presentation of a precise initial number can bias subsequent estimations.
- Legal Contexts: In legal proceedings, precise figures quoted for damages, compensation, or even proposed prison sentences can serve as anchors, influencing judges' or juries' final numerical judgments.

- Price Determination/Marketing: Retailers can leverage precise pricing (e.g., "19.95" instead of "20.00") to subtly influence consumer perceptions of value and willingness to pay, guiding their internal estimates of a "fair" price closer to the precise figure.

Bibliography

Dieter Rasch, L. R. Verdooren, and Jürgen Pilz. *Applied Statistics: Theory and Problem with R.* Willey, Oxford, 2019. ISBN 978-1-119-55152-2. doi: 10.1002/9781119551584.

Sarah Boslaugh, *Statistics in a Nutshell.* O'Reilly Media, Incorporated, 2012. ISBN 9781449361129.
https://books.google.de/books?id=s1llAQAACAAJ

David S. Moore, George P. McCabe, and Bruce A. Craig. *Introduction to the Practice of Statistics.* W.H. Freeman, New York, NY, USA, 7th edition, 2012.

Douglas C. Montgomery, Elizabeth A. Peck, and G. Geoffrey Vining. *Introduction to Linear Regression Analysis.* Wiley, Hoboken, NJ, 5th edition, 2012.

Andy Field, Jeremy Miles, and Zoë Field. Discovering Statistics Using R. Sage Publications, London, 1st edition, 2012.

Michael Schomaker, Christian Heumann, and Shalabh. *Introduction to Statistics and Data Analysis.*. Springer Cham, 2016. ISBN 978-3-319-46160-1.
https://books.google.de/books?id=s1llAQAACAAJ