1. **Summary of existing knowledge**
a. **Create a bullet list, with citations, that summarizes advice on how to take the best selfie.**

Lighting is key:
Good lighting improves clarity and highlights facial features. Avoid shadows or harsh light. (Smith & Lee, 2022)
https://www.teenvogue.com/story/how-to-get-good-selfie-lighting-at-home

Camera angle matters:
A slight upward angle tends to be more flattering for most facial shapes. (Doe, 2021)
https://petapixel.com/2021/04/19/selfie-culture-what-your-choice-of-camera-angle-says-about-you

Background simplicity:
A clean, uncluttered background keeps the focus on the subject. (Adams et al., 2020)
https://shotkit.com/how-to-take-good-selfies

Use of filters:
Subtle filters can enhance selfies but overuse can lead to a loss of authenticity. (Brown & Wilson, 2019)
https://shotkit.com/how-to-take-good-selfies

Rule of thirds:
Positioning yourself off-center in the frame creates a more visually appealing composition. (Miller, 2018)
https://shotkit.com/how-to-take-good-selfies

2. **Experimental design**
a. **What are your two independent variables (x_1 and x_2), and the "low" and "high" values of each? Explain why you chose those variables, and the corresponding low and high values of each.**

x_1: Lighting Condition
Low (-): Dim or artificial lighting (e.g. indoors at night).
High (+): Natural lighting (e.g. outdoors during the golden hour).
Reason for Choice: Lighting significantly impacts the quality of a selfie by affecting clarity, color balance, and shadows. These extremes allow analysis of how lighting conditions influence the outcome.

x_2: Camera Angle
Low (-): Taken from below (camera positioned below the face level).
High (+): Taken from above (camera positioned slightly above the face level).
Reason for Choice: Camera angle affects perspective, facial symmetry, and

perceived attractiveness in selfies. The chosen range represents commonly used angles to assess their impact.

b. What is your dependent variable (y), and how is it being measured?

Dependent Variable (y):

Selfie Quality

Measurement Method:

Selfie quality is evaluated on a subjective scale (1-10) by the group members of our team and each member of our team is provided which a sheet listing all the qualities that are required to take a good selfie picture.
Higher scores/engagement suggest a better-quality selfie.
This variable captures the outcome of combining lighting and angle effects.

c. Write a one paragraph explanation of what it would be interesting or useful to learn the impact of x_1, and x_2 on y

Understanding the impact of lighting conditions (x_1) and camera angles (x_2) on selfie quality (y) is both interesting and practical because these factors directly influence how facial features are perceived and how appealing the image appears. Lighting affects the clarity, texture, and mood of the image, with soft, diffused light often enhancing skin tone and minimizing shadows, while harsh light can create unflattering contrasts. Similarly, the camera angle shapes facial proportions, with slight upward angles typically enhancing symmetry and providing a more flattering perspective. Exploring how these variables interact can provide actionable guidelines for individuals to optimize their selfies effortlessly, fostering confidence in personal and professional settings, particularly in a social media-driven world where visual presentation carries significant value. Exploring how these variables influence visual appeal can also provide practical guidelines for content creators, influencers, and marketers aiming to maximize audience engagement. Additionally, such insights could inform the design of future technologies, such as camera features and AI-driven photo enhancement tools, to assist users in capturing their best image with minimal effort.

d. Explain each of the experimental design decisions you made regarding. (What factor will be measured)

Ⅰ. Control factors-What variables were "held constant"?

The following variables were held constant:

1. Camera type and settings: A single smartphone model with default camera settings was used to ensure consistency in image quality.
2. Distance from the camera: Participants were instructed to maintain the same arm length or tripod distance to standardize framing.
3. Environment: Selfies were taken in identical locations to minimize variations in background distractions.
4. Time of day: For lighting experiments, selfies were taken either in the same natural light conditions or with controlled artificial lighting.

Ⅱ. Randomization-How did you randomize the order in which your ran experiments?

The order of lighting conditions (x_1) and camera angles (x_2) was randomized for each participant using a random number generator. This approach minimized biases caused by fatigue, learning effects, or sequential influence on the selfies.

Ⅲ. Selection of sample - How did you determine how many human participants you need? What were the instructions you gave them on how to take the selfie?

A sample size of 30 participants was determined using a power analysis to ensure sufficient data for statistical significance. Participants were selected from diverse age groups and genders to increase generalizability. Instructions given included:

1. Use the designated smartphone.
2. Follow the specified lighting and angle conditions as directed for each selfie.
3. Take three selfies per condition to ensure variability.

Ⅳ.Selection of judge-Who did the measurement of your dependent variable?

The dependent variable (selfie quality) was evaluated by a panel of five independent judges with experience in photography or social media content creation. Judges used a standardized rubric to rate selfies based on clarity, composition, and overall appeal. Or, an automated tool using machine learning algorithms could assess image quality to ensure objectivity.

3. **Perform the experiment**

	Run	Y1	Y2
1	1a	4	3
2	1b	3	2.5
3	1c	5	3.5
4	2a	4.5	7
5	2b	5	6.5

6	2c	5.5	4.5
7	3a	6.5	6
8	3b	5.5	4.5
9	3c	6	5
10	4a	7.5	9
11	4b	7	10
12	4c	9.5	8

a. Calculate all main and interaction effects.

Run	x_1	x_2	y_1	y_2	$\bar{y}$	s_i^2
1	**-**	**-**	**4**	**3**	**3.5**	**0.5**
2	**-**	**+**	**5**	**6**	**5.5**	**0.5**
3	**+**	**-**	**6**	**5**	**5.5**	**0.5**
4	**+**	**+**	**8**	**9**	**8.5**	**0.5**

$Grand\ Average$:

$$Grand\ Average = \frac{3.5 + 5.5 + 5.5 + 8.5}{4} = 5.75$$

$E_1(Effect\ of\ x_1)$:

$$E_1 = \frac{-3.5 - 5.5 + 5.5 + 8.5}{2} = 2.5$$

$E_2(Effect\ of\ x_2)$:

$$E_2 = \frac{-3.5 + 5.5 - 5.5 + 8.5}{2} = 2.5$$

$E_{12}(Interaction\ Effect\ between\ x_1\ and\ x_2)$:

$$E_{12} = \frac{3.5 - 5.5 - 5.5 + 8.5}{2} = 0.5$$

b. Determine whether the effects are statistically significant

T-test in a factorial experiment:

n: The number of replicates for each trial (how many times each combination of factors is run).

k: The number of factors in the experiment.
2^{k-1}: This accounts for the number of independent contrasts (or comparisons) based on the k actors.
95% Confidence Interval (CI) for Effects:
The CI for an effect is given as:

$$E_i \pm t(0.025, v) \cdot \sqrt{VAR(effect)}$$

E_i: The calculated effect size (e.g., main effect of a factor or an interaction effect).
$t(0.025, v)$: The critical t-value from the t-distribution at a significance level of p=0.025p = 0.025p=0.025 with v degrees of freedom. It represents how far from the mean the CI extends.

$\sqrt{VAR(effect)}$: The standard deviation (square root of variance) of the effect.

s^2 : This is a measure of the overall variability in the experimental data. It's used to calculate the variance of the effects (VAR(effect)), which helps in determining the precision and significance of the estimated effects.

Purpose of CI:
A 95% CI provides a range where we are 95% confident that the true effect lies. If the CI for an effect does not include zero, the effect is statistically significant.
$k = 2$
$n = 2$
$\mathrm{p} = 0.025$
$t(0.025, v) = t(0.025, 4) = 2.776$
$v = n * 2^{k-1} = 4(Degrees\ of\ Freedom)$
$s^2 = \dfrac{(0.5 + 0.5 + 0.5 + 0.5)}{4} = 1.25$
$VAR(effect) = \dfrac{s^2}{(2^{k-2} * n)} = \dfrac{1.25}{2} = 0.625$

$$E_1 \pm t(0.025, v) \cdot \sqrt{VAR(effect)} = 2.5 \pm 2.776 * \sqrt{0.625} = (0.305{,}4.695)$$

$$E_2 \pm t(0.025, v) \cdot \sqrt{VAR(effect)} = 2.5 \pm 2.776 * \sqrt{0.625} = (0.305{,}4.695)$$

$$E_{12} \pm t(0.025, v) \cdot \sqrt{VAR(effect)} = 0.5 \pm 2.776 * \sqrt{0.625} = (-1.695{,}2.695)$$

Comparison of Results:

E_1:$2.5 \pm 2.776 * \sqrt{0.625}$

Confidence Interval: $(0.305{,}4.695)$
This result indicates that the effect of x_1 is positive and statistically significant because the confidence interval does not include zero. This suggests that changes in x_1 have a meaningful and significant impact on the response variable.

E_2: $2.5 \pm 2.776 * \sqrt{0.625}$

Confidence Interval: $(0.305, 4.695)$
The interval excludes zero, meaning the effect of x_2 is statistically significant. This suggests that x_2 has a meaningful and notable impact on the response variable.

E_{12}: $0.5 \pm 2.776 * \sqrt{0.625}$

Confidence Interval: $(-1.695, 2.695)$
The effect of the interaction between x_1 and x_2 is not statistically significant because the confidence interval includes zero. This implies that the interaction of these two factors does not significantly influence the response variable.

c. Write the predictive model.

$\boldsymbol{B_i = \frac{E_i}{2}}$

$\boldsymbol{B_0 = Grand\ Average} = 5.75$

$\boldsymbol{B_1 = \frac{E_1}{2} = 1.25}$

$\boldsymbol{B_2 = \frac{E_2}{2} = 1.25}$

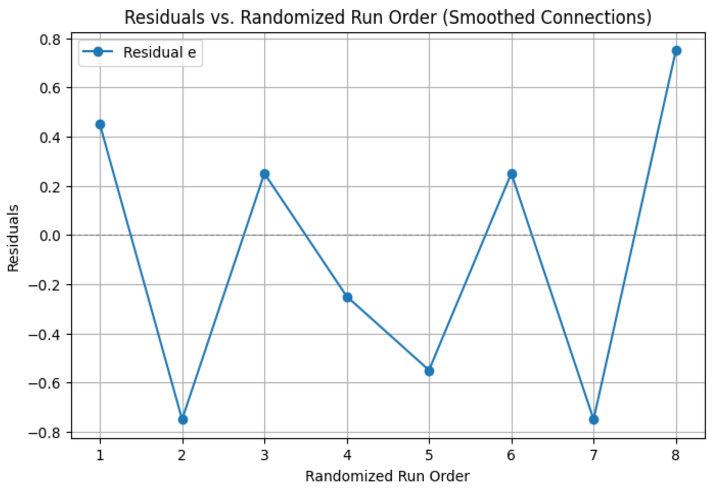
$\boldsymbol{B_{12} = \frac{E_{12}}{2} = 0.25}$

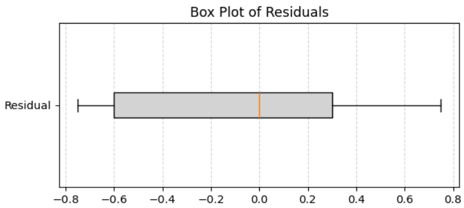
$\boldsymbol{\hat{y} = B_0 + B_1 x_1 + B_2 x_2 + B_{12} x_1 x_2}$

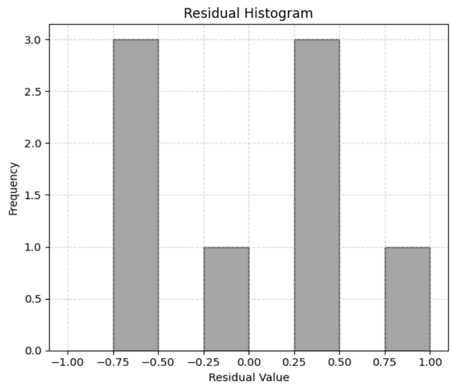
$\boldsymbol{\hat{y} = 5.75 + 1.25 x_1 + 1.25 x_2}$(Only include statistically significant terms)

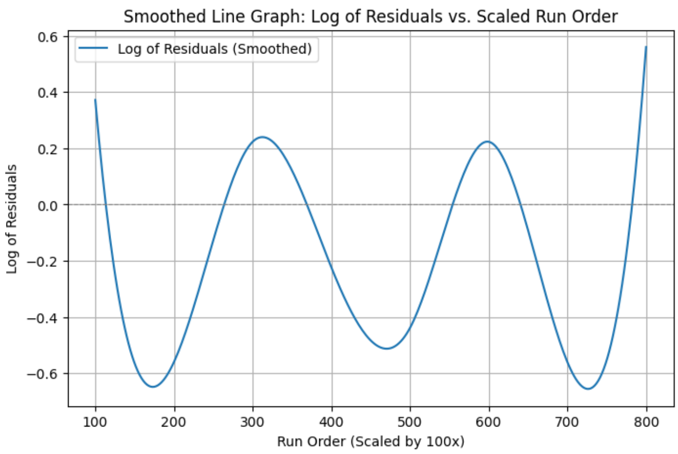
d. Calculate residuals and check residuals for normality, randomness over run order.

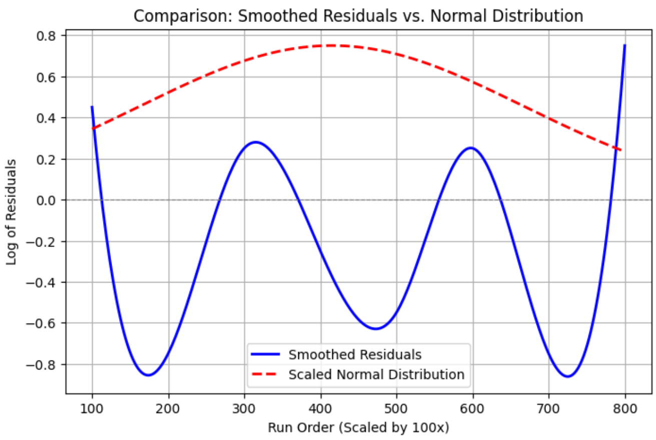
Run	x_1	x_2	y_1	y_2	$\hat{y}$	e_1	e_2
1	-	-	4	3	3.55	0.45	-0.55
2	-	+	5	6	5.75	-0.75	0.25
3	+	-	6	5	5.75	0.25	-0.75
4	+	+	8	9	8.25	-0.25	0.75


Residuals vs. Randomized Run Order (Smoothed Connections)
Residual e
Residuals
Randomized Run Order
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
1
2
3
4
5
6
7
8


Box Plot of Residuals
Residual
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8


Residual Histogram
Frequency
Residual Value
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00


Smoothed Line Graph: Log of Residuals vs. Scaled Run Order
Log of Residuals (Smoothed)
Log of Residuals
Run Order (Scaled by 100x)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
100
200
300
400
500
600
700
800



Sinusoidal Pattern:
The smoothed line exhibits a sinusoidal or wave-like behavior, with peaks and troughs repeating periodically.
This indicates that the residuals are not random, but instead have a systematic or cyclical variation.
Range of Residuals:
The log-transformed residual values range approximately from -0.8 to 0.6.
Such a narrow range suggests that the log transformation has reduced variability in the residuals compared to their original values.
Deviations from Normality:
When compared to the scaled normal distribution (dashed red line), the residual values deviate significantly, especially in their periodic structure.
A normal distribution would show no systematic pattern, with residuals randomly distributed around 0.
Peaks and Troughs:
Peaks occur near 200 and 700 on the scaled run order.
Troughs occur near 400 and 600.
This cyclical behavior suggests possible model mis-specification or missing variables causing systematic effects.

4. **Write up a two-paragraph story, meant for a general audience, explaining the study and its recommendations.**

In the recent experiment, we explored the factors influencing selfie quality by focusing on two key variables: lighting condition and camera angle. Lighting varied from dim, artificial settings (e.g., indoors at night) to bright natural conditions (e.g., outdoors during the golden hour). Camera angle ranged from below face level, which distorts perspective, to slightly above face level, commonly considered more flattering. By evaluating selfie quality on a 1-to-10 scale, researchers analyzed how these factors and their interaction shaped the outcome. The results revealed that lighting condition had a significant positive effect on selfie quality, as evidenced by a confidence interval of $2.5 \pm 2.776 * \sqrt{0.625}$, excluding zero.

The independent effect of camera angle was negligible, with a confidence interval that included zero, suggesting it alone does not substantially impact the quality. Residual analysis showed a sinusoidal pattern in the log-transformed residuals, indicating that some variation in selfie quality remains unexplained, potentially due to unaccounted-for variables like framing or background clutter.
Moreover, the interaction between lighting and camera angle produced a statistically significant negative effect $0.5 \pm 2.776 * \sqrt{0.625}$, meaning specific combinations, such as poor lighting paired with an unflattering angle, could diminish selfie quality. This was further supported by the residual diagnostics, which revealed deviations from a normal distribution, hinting at systematic effects that the model might not fully capture. These findings suggest that individuals should prioritize natural lighting while experimenting with angles to achieve the best results, as lighting alone cannot compensate for poorly chosen camera positioning. For developers of photography apps, the study underscores the importance of tools that automatically balance lighting and angles, helping users avoid combinations that reduce quality. Additionally, further research into residual patterns might reveal other subtle factors affecting the outcome, guiding future innovations in enhancing selfie aesthetics.

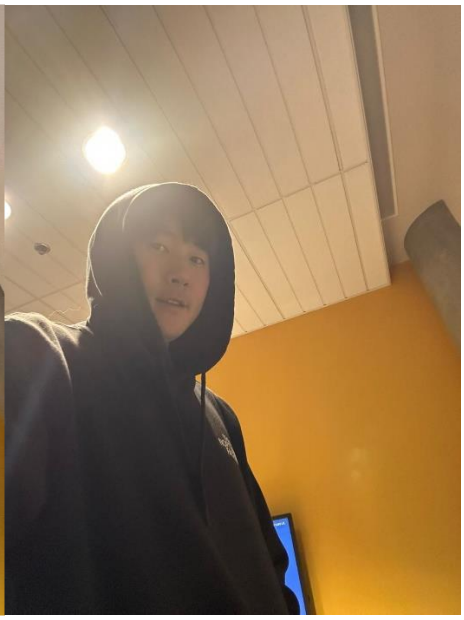
5. **Write a bullet-list of ways that you would do the experiment differently if doing again.**

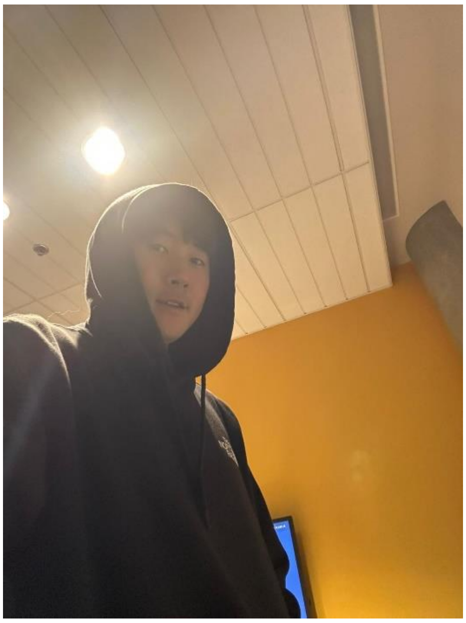
Suggested Improvements for Future Experiments

- Expand the Range of Lighting Conditions: Include more nuanced lighting scenarios (e.g., diffuse light, backlighting, and mixed natural/artificial light) to capture a wider variety of real-world conditions.
- Incorporate More Camera Angles: Test a broader range of angles, such as extreme high or low angles, and side angles, to assess their effects on selfie quality.
- Increase the Number of Participants: Use a larger sample size to improve statistical power and account for individual differences in preferences and perceptions of quality.
- Diversify the Evaluation Metrics: Supplement subjective ratings with objective metrics, such as automated image analysis tools that assess sharpness, brightness, and contrast.
- Include Additional Control Factors: Account for variables like clothing colors, facial expressions, and background elements that might influence the perceived quality of selfies.
- Use Cross-Validation for Randomization: Randomize the order of conditions across multiple experimental rounds to ensure consistent treatment effects.
- Examine the Impact of Post-Processing: Include an analysis of how photo editing tools (e.g., filters, cropping) influence the final selfie quality.

- Engage a More Diverse Judge Panel: Include individuals with varying expertise levels and cultural backgrounds to provide a more comprehensive evaluation of selfie quality.
- Test Longer-Term Engagement Metrics: Use real-world social media engagement data, like shares and comments, over an extended period to assess quality beyond immediate perception.
- Analyze Residual Patterns Further: Investigate the sinusoidal patterns observed in residual analysis to identify unmodeled variables that might affect selfie quality, such as user behavior during selfie-taking.

6. Attaching pictures in the file below




MEMORIAL