1   Proportional Reasoning Across Formats

2   Maanav Choudhary[1]

3   [1] Rutgers University

4   Author Note

₁₀   Abstract

₁₁ Comparing proportions is a fundamental skill that even infants seem to possess, albeit at a

₁₂ rudimentary level. This study explores how well people compare proportions when

₁₃ presented in different visual formats, aiming to deepen our understanding of this cognitive

₁₄ process. Specifically, we examined whether average performance varies across format types,

₁₅ whether numerator congruency status influences performance, and whether an interaction

₁₆ exists between format type and numerator congruency. Here, we show that both format

₁₇ type and numerator congruency significantly impact performance, with notable interactions

₁₈ between these factors. These findings reveal that certain formats and congruency

₁₉ conditions facilitate more accurate comparisons, advancing previous knowledge by

₂₀ highlighting the importance of presentation design in proportion reasoning. This work

₂₁ contributes to a broader understanding of cognitive processing in numerical tasks and has

₂₂ implications for designing more effective educational and communicative tools.

₂₃   *Keywords:* keywords

₂₄   Word count: X

₂₅   Proportional Reasoning Across Formats

₂₆   **Methods**

₂₇   Here we outline how the data was collected, as well as information on the participants

₂₈ and materials used.

₂₉ **Participants**

₃₀   A total of 99 adults participated in this study.

₃₁ **Material**

₃₂   The images provided to the participants had a mix of blue and orange. Some images

₃₃ had these represented as blobs, and some as dots. There were four conditions for the mix of
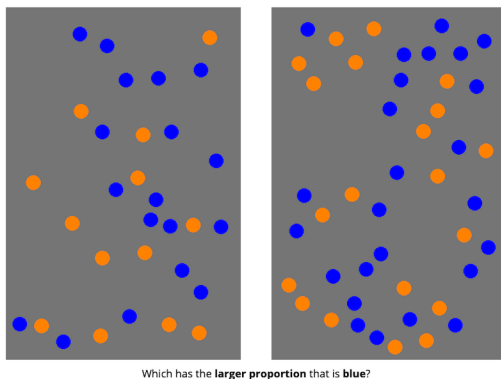
₃₄ blue and orange.

Which has the **larger proportion** that is **blue**?

Figure 1. Example of an image given to participants.

₃₅ **Procedure**

₃₆   First, participants were introduced to a story about a magic ball and that the

₃₇ outcome (i.e., blue or orange) depended on the proportions. Then they were asked to

₃₈ compare the proportions of different images, for example the ones shown in Figure 1

⁣₃₉ In other words, participants were shown to images of the same kind at the same time
⁣₄₀ and asked to decide which had a higher proportion of the shape (or dots) colored in blue.

Data analysis

⁣₄₂ We used R (Version 4.4.1; R Core Team, 2024) and the R-packages *dplyr* (Version
⁣₄₃ 1.1.4; Wickham, François, Henry, Müller, & Vaughan, 2023), *ggdist* (Version 3.3.2; Kay,
⁣₄₄ 2024), *ggplot2* (Version 3.5.1; Wickham, 2016), *papaja* (Version 0.1.3; Aust & Barth, 2024),
⁣₄₅ *tibble* (Version 3.2.1; Müller & Wickham, 2023), *tidyr* (Version 1.3.1; Wickham, Vaughan,
⁣₄₆ & Girlich, 2024) and *tidyverse* (Version 2.0.0; Wickham et al., 2019) for all our analyses.

Results

1. Does average performance vary across format types, ignoring all other aspects of the
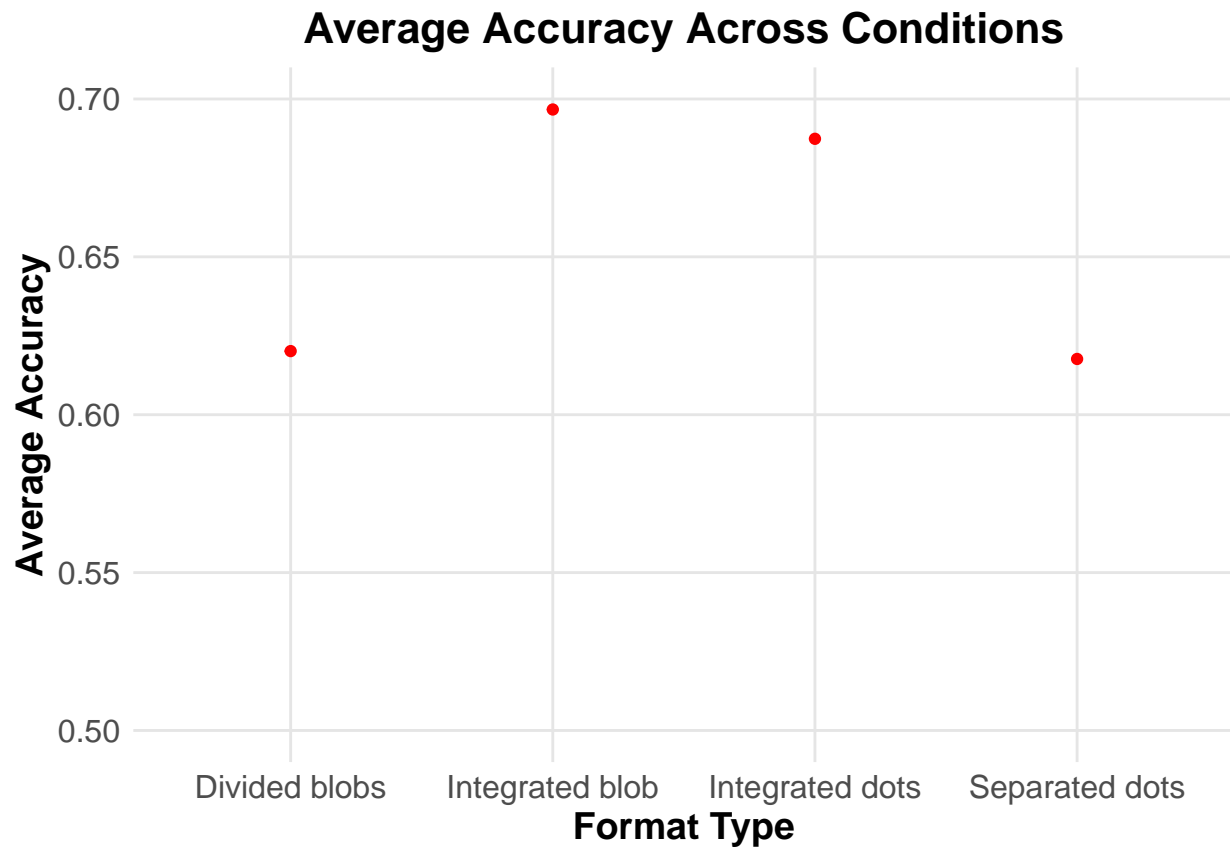
stimuli?

Figure 2. Average Performance.

Figure 2 shows the average accuracry for all format types, showing that the

integrated blobs and integrated dots both showed a higher average than the separated

blobs and dots, the averages of the priors being close to each other(.01 apart) and the

averages of the latter being close to each other (.003 apart)
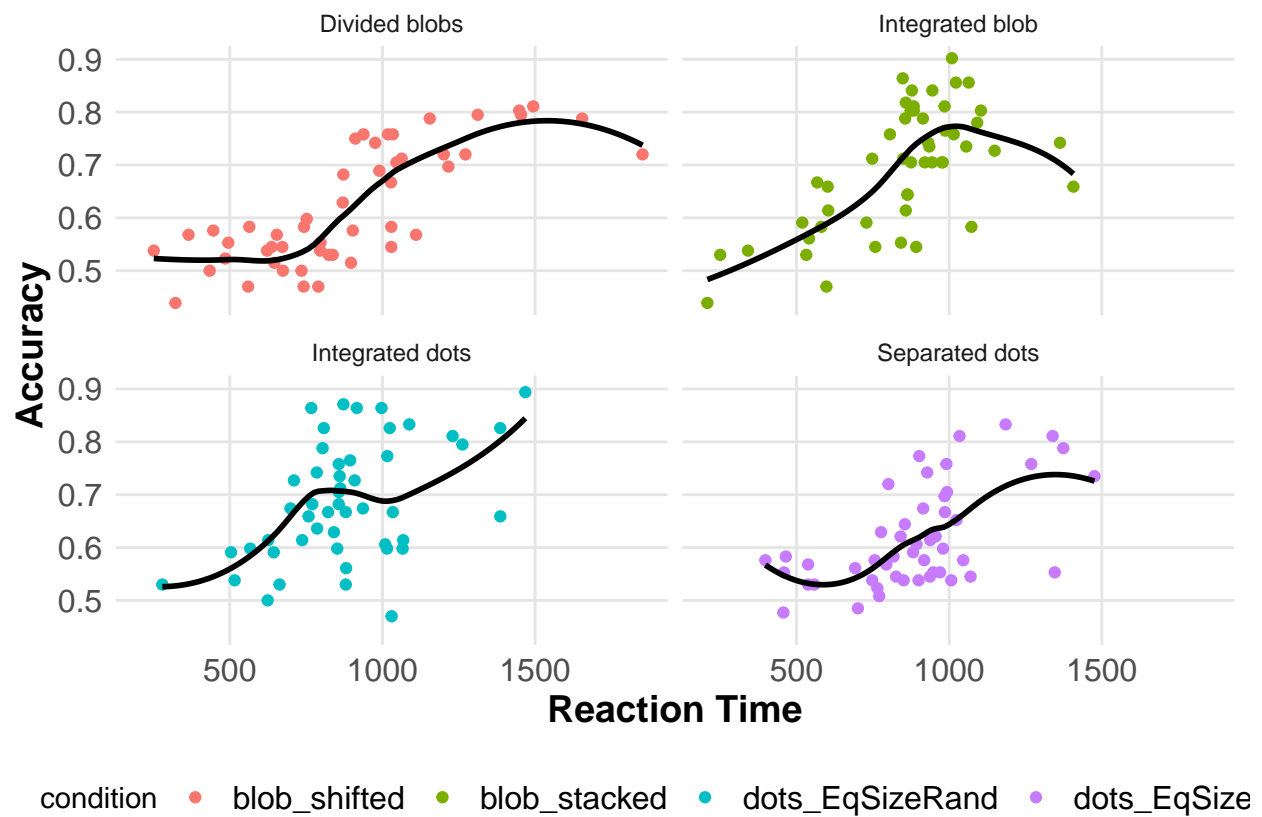
2. How are reaction time and accuracy related?

Figure 3. Reaction Time

⁵⁵ Figure 3 outlines that across all format types, as the reaction time becomes longer,

⁵⁶ accuracy rises. The lines of best fit show that in most conditions this increase drops off

⁵⁷ after a certain point.

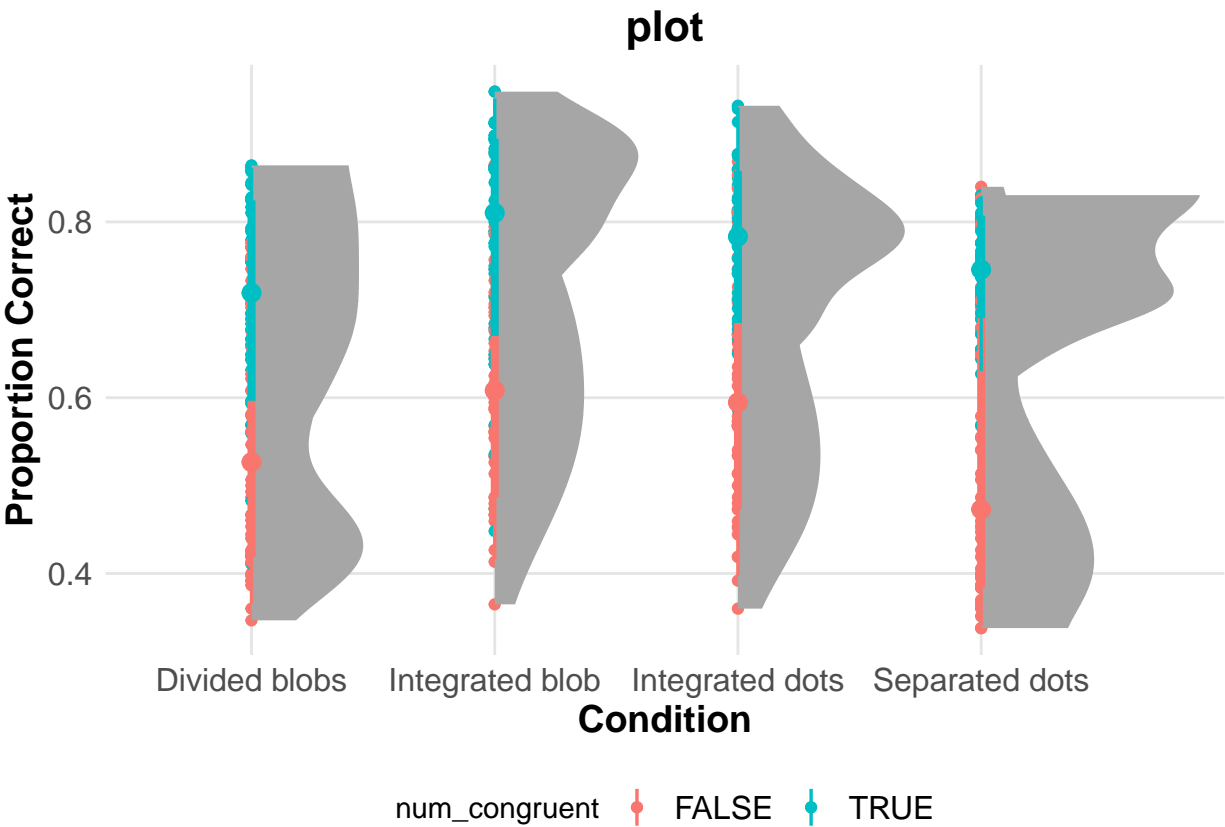58   3. How does numerator congruency interact with format type?

Figure 4. Numerator Congruency and Correctness

59   The figure 4 above shows that, for all conditions, when the numerators are congruent

60   the proportion of correct responses is higher.

Discussion

Interpretation

63   1. Does average performance vary across format type?

64   The data shows that the average performance does indeed vary across format type.

65   The data indicates that when blobs and dots are integrated, the performance is higher than

when they are separated. The performance difference between blobs and dots when they are both separated or integrated is not very large (.68 and .69 for integrated and .617 and .62 for separated). The higher averages for integrated data vs separated that this data shows indicates that humans may deal better with integrated visuals than separated when gauging quantity.

2. Does average performance vary across numerator congruency status?

The average performance is higher on when the numerators are congruent then when they are not.

3. Does numerator congruency vary across format type? (i.e., is there an interaction)

Numerator concurrency seems vary slightly across format type. Both of the dot cases appear to have less numerator concurrency than the blob conditions.

Conclusion

1. One of the hardest parts of the assignment was interpreting how the data and graphs related to the research objectives, and trying to turn the visual information into something written and coherent.

2. The most satisfying and enjoyable part of the assignment was figuring out how to insert the images, and playing around with the format to make it visually appealing.

References

Aust, F., & Barth, M. (2024). *papaja: Prepare reproducible APA journal articles with R Markdown.* https://doi.org/10.32614/CRAN.package.papaja

Kay, M. (2024). ggdist: Visualizations of distributions and uncertainty in the grammar of graphics. *IEEE Transactions on Visualization and Computer Graphics*, *30*(1), 414–424. https://doi.org/10.1109/TVCG.2023.3327195

Müller, K., & Wickham, H. (2023). *Tibble: Simple data frames.* Retrieved from https://CRAN.R-project.org/package=tibble

R Core Team. (2024). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A grammar of data manipulation.* Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., Vaughan, D., & Girlich, M. (2024). *Tidyr: Tidy messy data.* Retrieved from https://CRAN.R-project.org/package=tidyr