

The Aesthetics of Knowledge Consumption

Does aesthetics matter in popular science media?

Leoson Hoay

Masters in Computational Social Science, Social Sciences Division, University of Chicago

Introduction

Aesthetics and the perceived value of objects are inextricably linked. All human communicative acts, including science communication, are not spared from the eye of the beholder. This study aims to examine the impact of aesthetics on the modern consumption of scientific knowledge, by focusing on a popular medium of information consumption today – the internet.

New media and psychology studies have measured aesthetic perception in the experimental setting - however, resources are not always available to conduct large-scale studies using the experimental paradigm.

This study borrows from HCI and UI studies **in applying computational aesthetic measures to examine science news websites**. These measures are then examined for their association with key performance metrics of these media outlets.

Aims & Hypotheses

Are computational aesthetic measures of website interfaces associated with better web metrics on science news websites?

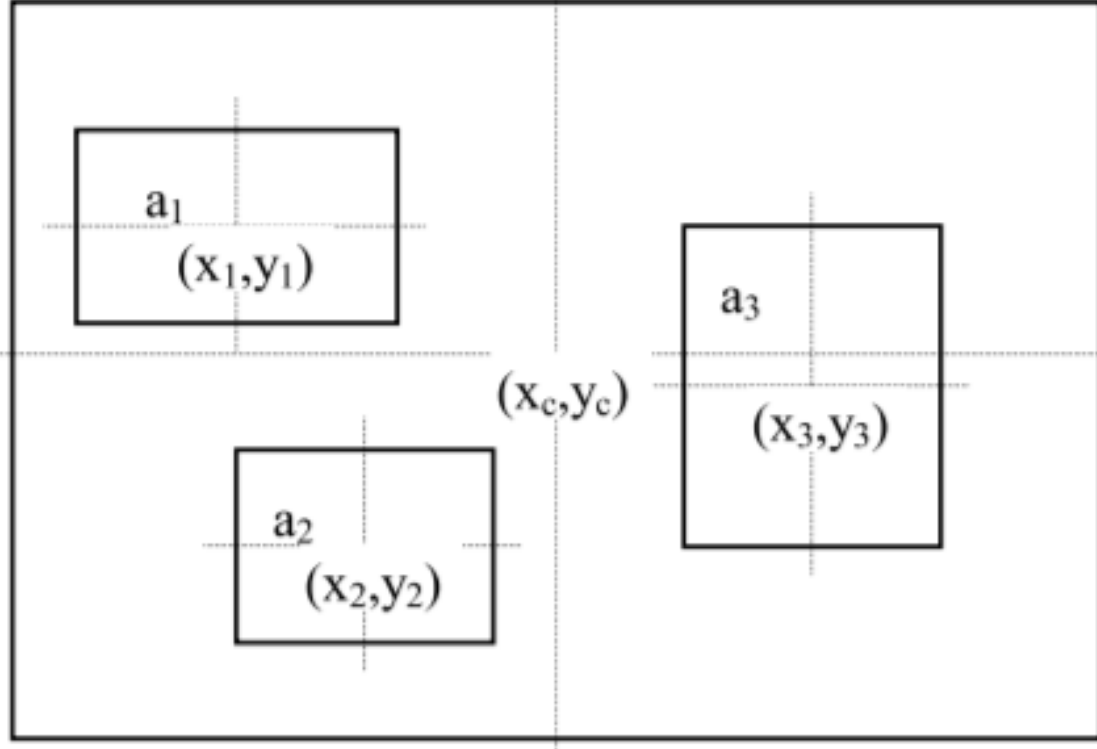
This study examines two foundational website metrics - average site visits for the month, and average user linger time

H1: There is a significant relationship between aesthetic metrics of science web articles and the **readership of the media outlet they are from**.

H2: : There is a significant relationship between aesthetic metrics of science web articles and **how long users spend on the website** on average.

Aesthetic Measures

Figure 1: Model of Screen Equilibrium by Ngo et. al, 2000



TEXT AESTHETICS

The aesthetics of article text is operationalized as stylistic consistency, and implemented here as the **document similarity** of each article (or the mean similarity between media outlets at the top level) with reference to a corpus of its other articles from the same media outlet. The *gensim** python library was employed for this purpose, and the articles scraped using the *boilerpipe** python wrapper.

*Credit to the creators of these modules – Hidayat, Kohlschutter, Rehurek, and others. Example boilerpipe and PhantomJS code is available on my github page (bottom left).

LAYOUT AESTHETICS

Six computational aesthetic measures, drawn from HCI and UI metrics, are considered in this study. The algorithms were implemented with the help of *OpenCV**, an open source computer vision software library. The web article images were scraped with the help of the *PhantomJS** scriptable.

Layout Aesthetic Metrics

Kolmogorov Complexity	$M_D(r) = 1 - avg_{i < j, r} \{NID(i, j)\}$
Shannon Entropy	$M_I(r) = \frac{I(X_i, Y_r)}{H(X_i)}$
Colorfulness	Measure of pixel color differentials
Edge Density	Measure of edge color gradients
Screen Equilibrium	$(x_c, y_c) = \left(\frac{\sum_i a_i x_i}{\sum_i a_i}, \frac{\sum_i a_i y_i}{\sum_i a_i} \right)$
Screen Sequence	$sgm = (p, d)$

Descriptives and Aesthetic Metrics

Media Outlets	Articles From Each Outlet	Total Data Points (n)
25	30	750

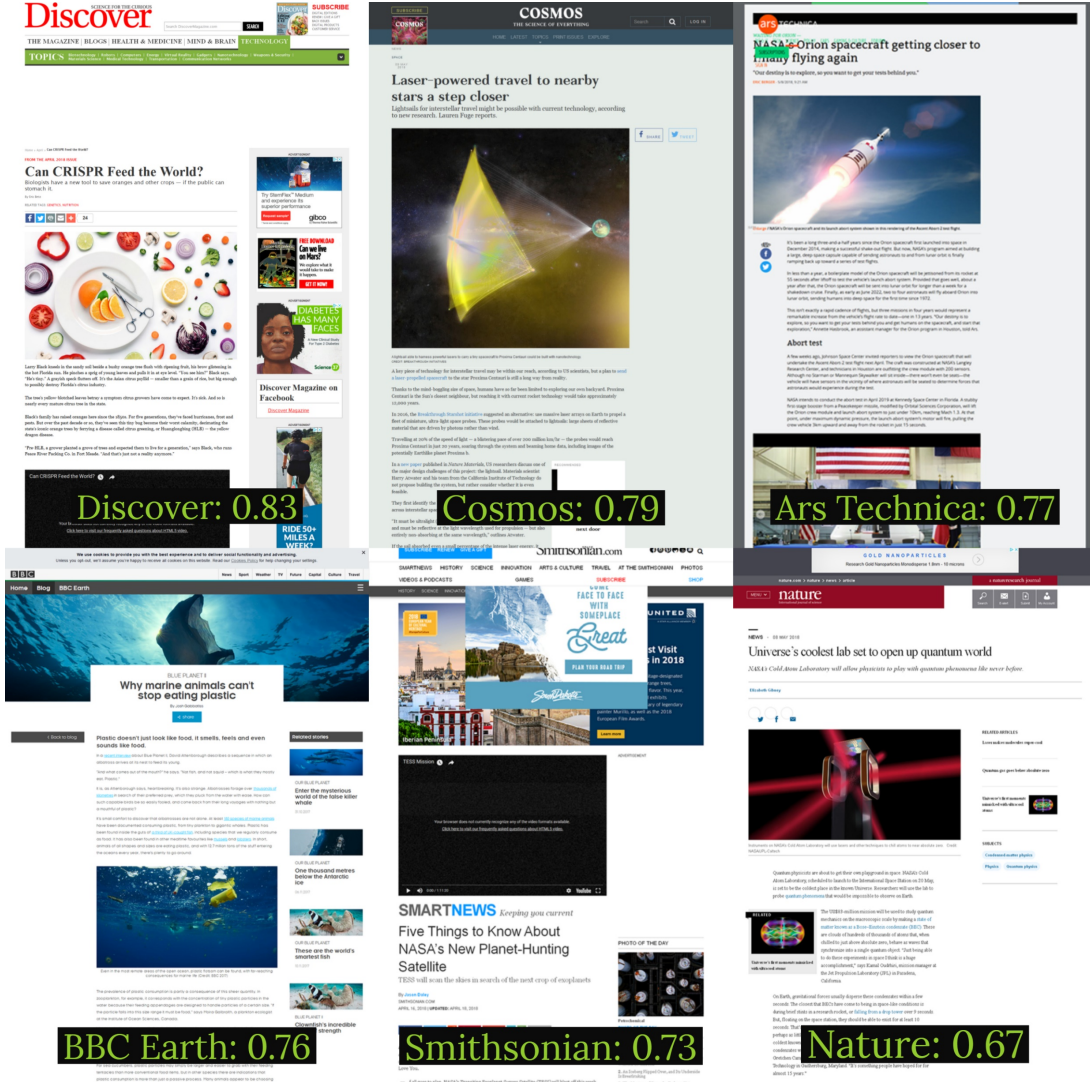
Table 1: Document Consistency Measures for 5 Outlets

Media Outlet	Average Document Consistency
National Geographic	0.76
Wired	0.63
Live Science	0.73
Science Magazine	0.54
Scientific American	0.71

Table 2: Screen Equilibrium Measures for 5 Outlets

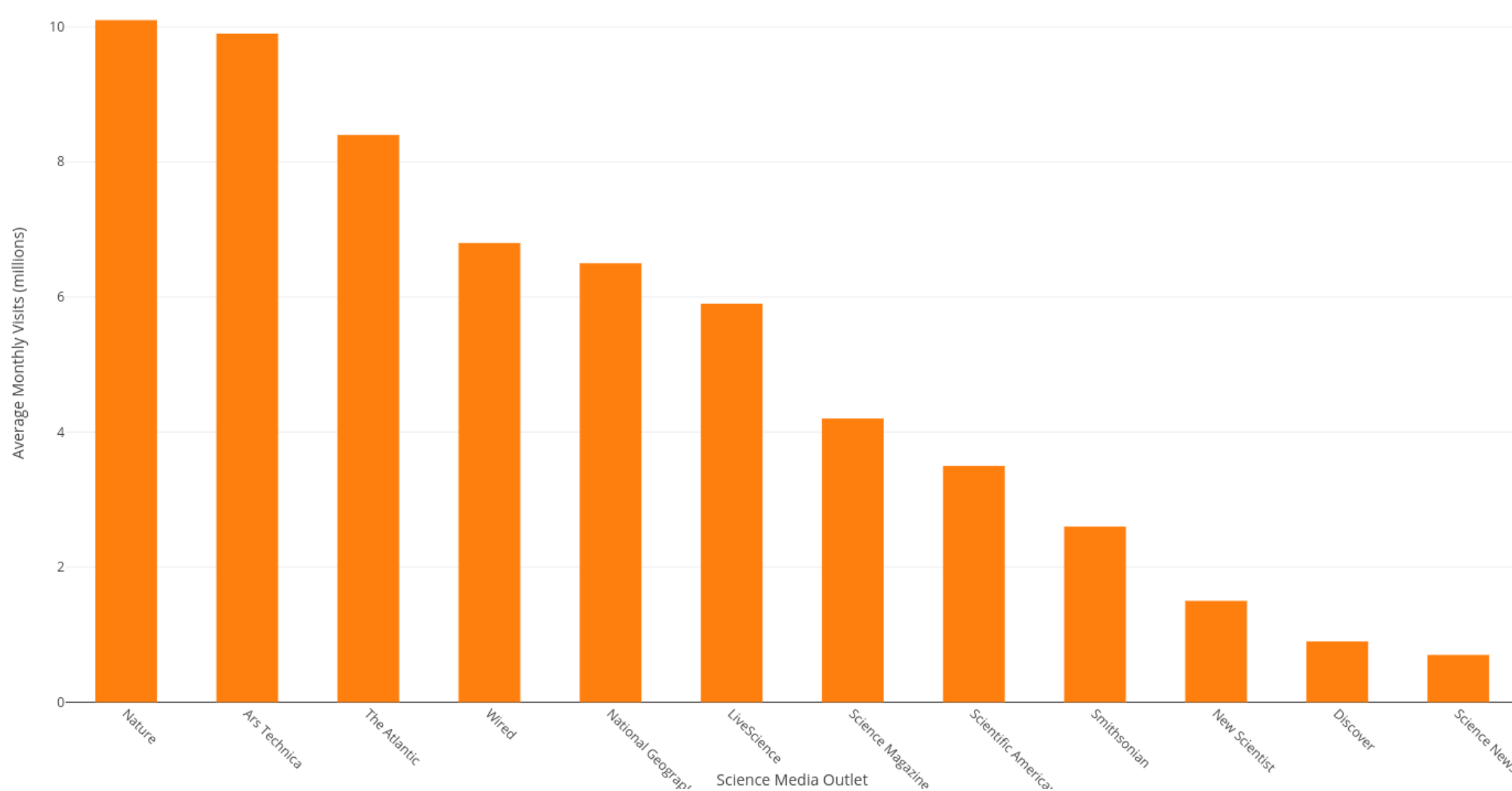
Media Outlet	Average Screen Equilibrium
National Geographic	0.79
Wired	0.86
Live Science	0.76
Science Magazine	0.76
Scientific American	0.82

Figure 2: Colorfulness Score for 6 Outlets

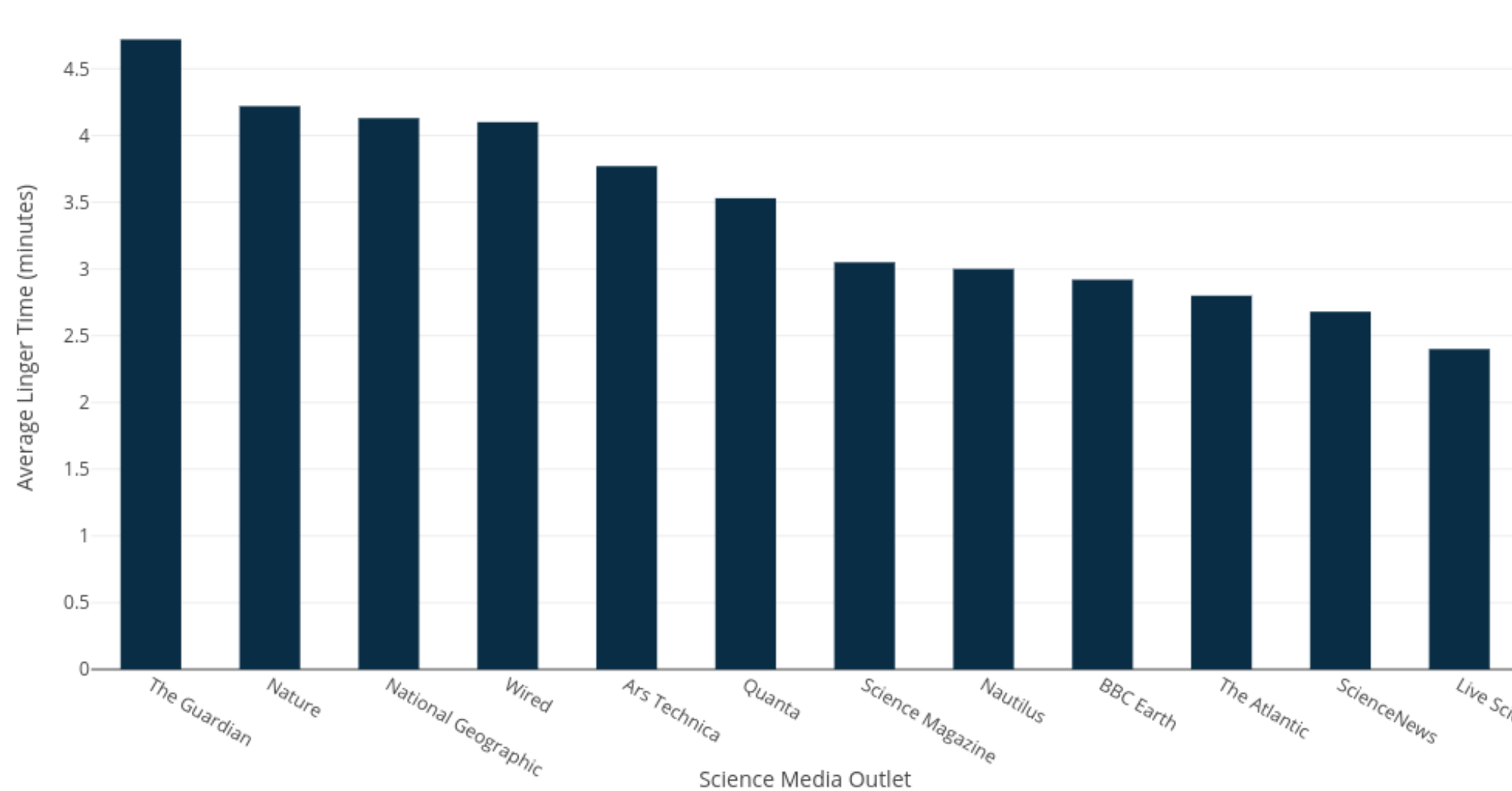


Science Media Outlets Considered				
Scientific American	The Guardian	Smithsonian	Undark	The Atlantic
New Scientist	Nature	BBC Earth	Nautilus	Science Mag
Live Science	Discover	NatGeo	Cosmos Magazine	Quanta
Science News	Wired	Ars Technica	Atlas Obscura	Gizmodo
Space	Futurism	Science Daily	SciTechDaily	Phys.Org

Top 12 Monthly Site Visits (April 2018)

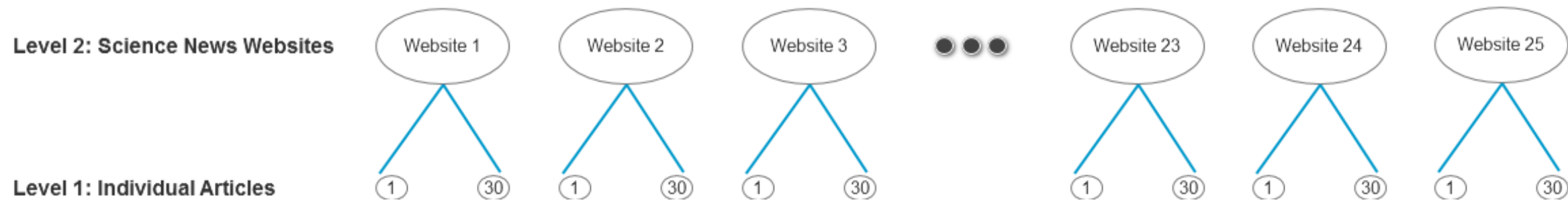


Top 12 Average User Linger Times (April 2018)



Multi-level Model

We build a multi-level model where the first level represents the individual science articles, and the second level represents the various media outlets.



Article Level:

$$V_{ij} = \beta_{0j} + \beta_{1j}c + \beta_{2j}\omega + \beta_{3j}\alpha + \beta_{4j}x_1 + \beta_{5j}x_2 + \beta_{6j}x_2^2 + \beta_{7j}x_3 + \beta_{8j}x_4 + \epsilon_{ij}$$

Media Outlet Level:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + v_{0j}$$

For article (i = 1 ... 30) and media outlet (j = 1 ... 25). The variables at the article level represent the aesthetic measures, and Z represents the number of years since a specific website j was inaugurated, as of April 2018.

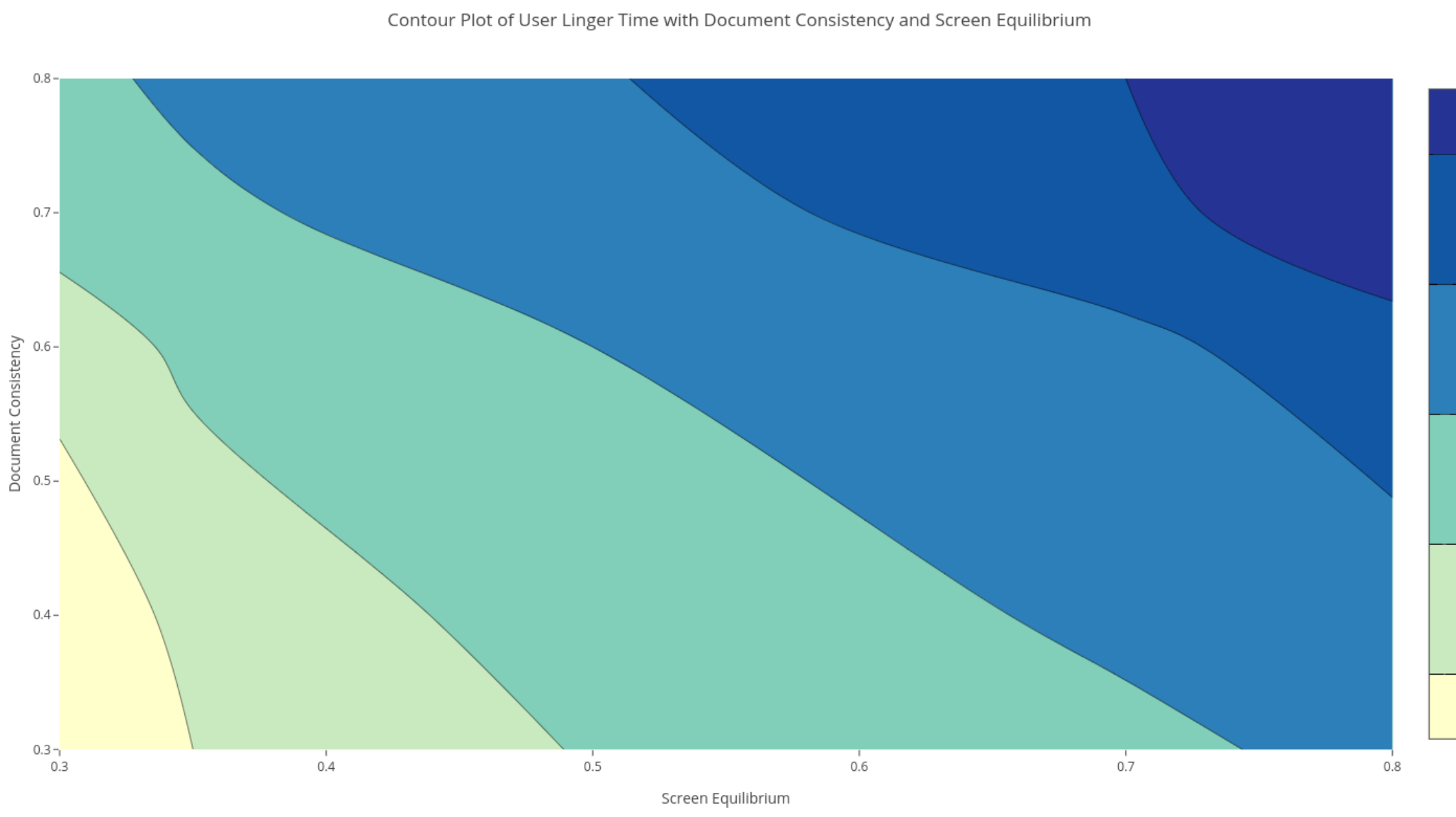
Both readership and linger time run on the similarly specified models. LDA on the aesthetic variables reveal that Edge Density and Kolmogorov Complexity have negligible loadings and contributions to overall variance, and hence were dropped from the model.

Results

Only the site visits model has been fully calibrated, an unfortunately it does not reveal any significant association between aesthetic variables and monthly site visits. This could be due to various factors, including extraneous influences and errors in site visit estimation.

Fixed Part	Coefficient (s.e.)
Intercept	0.12 (0.20)
Document Consistency	0.51 (0.33)
Shannon Entropy	0.76 (0.48)
Colorfulness	0.31 (0.22)
Colorfulness2	1.10 (0.97)
Screen Equilibrium	1.73 (0.90)
Screen Sequence	1.34 (0.78)
Time since inauguration	0.31 (0.08)*
Random Part (omitted from table)	

A subset of the data has been run through a preliminary model for user linger time, which displays some potential in identifying significant relationships between document consistency and screen equilibrium with average user linger time.



Conclusion

- Aesthetic variables are unable to significantly predict visits
- Aesthetic variables have potential to significantly predict linger time to a marginal degree.
- Variations in aesthetic variables between each individual article are minimal, hence allowing for the specification of an “all-contextual, no-individual” model

Limitations

- Independent variables were measured using estimated method (SEMRush), with substantial error intervals
- Given more resources and time, could obtain organic data from media outlets
- Lack of page specific metrics prevents creation of model with more granularity (individual page hits/buzz)

References

Altobelli, A., & Lin, Y. (2011). Investigating Effects of Screen Layout Elements on Interface and creen Design Aesthetics [Research article]. Advances in Human-Computer Interaction. <https://doi.org/10.1155/2011/659758>

Birkhoff, G. D. (1933). Aesthetic Measure: Cambridge, MA and London, England: Harvard University Press. <https://doi.org/10.4159/harvard.9780674734470>

Moshagen, M., & Thielsch, M. T. (2013). A short version of the visual aesthetics of websites inventory. Behaviour & Information Technology, 32 (12), 1305-1311. doi: 10.1080/0144929X.2012.694910

Ngo, D. C. L., Samsudin, A., & Abdullah, R. (2000). Aesthetic Measures for Assessing Graphic Screens. Journal of Information Science and Engineering, 20.

Norman, D. (2013). The Design of Everyday Things (Revised, Expanded Edition). Basic Books.

Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, J., & Gajos, K. Z. (2013). Predicting users’ first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (p. 2049). Paris, France: ACM Press. <https://doi.org/10.1145/2470654.2481281>

Rigau, J., Feixas, M., & Sbert, M. (n.d.). Conceptualizing Birkhoff’s Aesthetic Measure Using Shannon Entropy and Kolmogorov Complexity. Computational Aesthetics in Graphics, Visualization, and Imaging.

Wittgenstein, L. (1973). Philosophical Investigations (3rd Edition). Pearson.

Acknowledgements

MACS_lecturers = [“Benjamin Soltoff”, “Rick Evans”]
MACS_preceptors = [“Joshua Mausolf”, “Ryan Hughes”]
from uchicago_student_lists import MACSS_2019
for student in MACSS_2019:
 thank(student)
for lecturer in MACS_lecturers:
 thank(lecturer)
for preceptor in MACS_preceptors:
 thank(preceptor)