

Title: Effects of Frequency Distribution on Linear Menu Performance
Authors: Wanyu Liu, Gilles Bailly, Andrew Howes
DOI: <https://doi.org/10.1145/3025453.3025707>

Motivation: Before the first lecture I had not given much thought about the possible implications of menu design and was curious about the research. Initially I only intended to skim through the paper, but because of the sheer number of topics that I either strongly disagreed on or greatly questioned during lecture 2, I wanted to dig deeper.

Contribution Type: I would categorize this paper as an empirical study because it discussed and trialed a phenomenon that's been widely accepted in the scientific community (Zipf's law). Additionally, I would classify it as a method contribution since it actually uses zipfian distributions which are often avoided due to their complexity in similar studies.

Scope: This paper trials different distributions for menus and how they impact a human user's ability to locate a desired menu item.

Approach: The trials utilize three different distribution types to test how quickly a correct menu, and subsequent menu item is selected by the human user. A uniform distribution was used to establish a baseline, and then two zipfian distributions with an s-parameter of 1 and 2 respectively as comparisons. The base assumption was that menus that follow a zipfian distribution are more efficient to use due to frequently used items occurring earlier.

Evidence: 24 participants collectively performed around 16K trials which is a reasonable number, but based on how these people were selected, many of them were likely quite similar which could add some form of hidden bias to the results. With that said, even at this scale the study was able to derive that the location of a menu is indeed faster to find with a zipfian distribution compared to a uniform distribution, but unintuitively, the reverse applies to finding menu items.

Generalizability & Alternative Applications: With how universal menus are in software and how universal software itself is, I'm not sure how this could be "broadened" any further.

Learning outcome & Discussion: While I certainly learned some new things and gained perspective on the matter, I'm still not convinced of the methodology. The author's use the time it takes for a user to locate a menu item to assess a distribution's performance, which I am not sure makes sense or is particularly useful for any real world application. Is saving a fraction of a second on average truly worth the potentially infinite amount of frustration that a "living" dynamic menu would cause? This type of design makes me think that the author's might be "reverse anthropomorphizing" humans by assuming that we work the same as machines. I claim this because the idea sounds perfect for an algorithm but I wouldn't wish it on my worst enemy. Regardless, the contradictory findings were quite eye opening how one type of methodology can work well for one type of menu, but be completely detrimental to another, even when both menus are adjacent.

Title: Modeling Sub-Document Attention Using Viewport Time

Authors: Max Grusky, Jeiran Jahani, Josh Schwartz, Dan Valente, Yoav Artzi, Mor Naaman

DOI: <https://doi.org/10.1145/3025453.3025916>

Motivation: As a developer in the web-space, it has always bothered me that data collection is done with the notion of quantity over quality. This often leads to relatively simple websites with abysmal performance due to the sheer amount of embedded trackers in them.

Contribution type: In my opinion, this paper falls under the category of “Theory and modelling”. My justification for this choice is because the paper does not necessarily produce a revolutionary end result, but is able to come very close to the results of cutting-edge eye-tracking studies that are far more expensive and difficult to conduct.

Scope: The paper leverages a relatively simplistic metric called “viewport time” to measure the reading rate of consumers on a website with long-form articles. This aforementioned reading rate is derived from how long certain HTML elements on the website spend on the user’s display. The reason why this is significant was covered in the previous paragraph.

Approach: The paper’s main contributions are three increasingly complex models that all strive to estimate the user’s reading rate by using viewport time. The naive baseline model (UAM) assumes that the reader’s attention is divided uniformly across each HTML element in the document. In the second model (UVAM), the user’s estimated attention is uniformly distributed across all currently visible HTML elements as well as assigning an even probability for any pixel on the screen to be viewed. The third and final model (GVAM) works very similarly to the UVAM model, but the estimated viewport attention equation uses the standard deviation and mean value from a previous study on reading rates.

Evidence: To validate the model estimated reading rates, the results of an empiric, cross-language study was used. Based on the presented comparisons, both the UVAM and GVAM models significantly outperformed the baseline UAM model, and the more complex GVAM model was able to gain an advantage over the UVAM model. The GVAM model ended up with one clear outlier language (Spanish), and the underlying reason for its performance was never identified. The dataset for the models is certainly large enough, but perhaps the authors were too strict when pruning “unacceptable users”. If all the remaining users act more or less the same way, that would make the paper’s results less significant since they would only be able to estimate a certain type of user’s behaviour.

Alternative applications & Learning outcome: I started wondering whether this type of “effective low tech” solution could be used to detect and reduce the spread of fake news and clickbait by verifying whether the important sections of the article were read at a similar rate to the title and early paragraphs. It is a little depressing that this type of technology seems fixated on finding new ways of exploiting human psychology to keep readers engaged for longer, rather than trying to improve human discourse, but perhaps there’s room for that in another paper.