1 TBD

2 Matthew J. C. Crump[1], Walter Lai[1], & Nicholaus Brosowsky[1]

3 [1] Brooklyn College of the City University of New York

4 Author Note

5 In this draft, listed authorship order simply indicates who is participating in the project

6 Correspondence concerning this article should be addressed to Matthew J. C. Crump,

7 Brooklyn College of CUNY, 2900 Bedford Avenue, Brooklyn, NY, 11210. E-mail:

8 mcrump@brooklyn.cuny.edu

Abstract

Enter abstract here. Each new line herein must be indented, like this line.

Keywords: keywords

Word count: X

₁₃ TBD

₁₄ Hicks Law refers to the noticeable slowing of responses in choice reaction experiments

₁₅ when the possible stimuli are less predictable.

₁₆ Hicks in his experiment, "On the rate of gain of information", tested his reaction speed

₁₇ on a choice reaction test. During each trial he had to determine which of n lights was turned

₁₈ on. He manipulated the set size n, and found that the greater the number of alternatives the

₁₉ slower his response. His data also suggested that reaction speed is linearly related to the

₂₀ predictability of responses; Trials in which certain responses were disproportionately more

₂₁ likely ellicited faster response times than those in which each response was equally likely.

₂₂ Hick's law at larger set sizes however does not have consistent experimental support.

₂₃ Conrad (1962) in an experiment which required subjects to name 320 nonsense

₂₄ syllables within trials of variable number of distinct syllables (number of alternatives) found

₂₅ that the response time was linearly correlated with the logarithm of number of alternatives.

₂₆ However, Pierce and Karlin in 1957 in an experiment which required subjects to read pages

₂₇ of words as quickly as possible, reported no differences in response times between trials with

₂₈ set sizes that ranged from 4 to 256 words.

₂₉ Proctor and Schneider in a review of studies testing Hick's Law attributes the

₃₀ inconsistency in large set size experiments to differences in skill level or practice between

₃₁ subjects and the arbitrariness of Stimulus Response coding/mappings when creating more

₃₂ number of alternatives.

₃₃ Our experiment sought to test Hick's law on typing speeds.

₃₄ In this study, we used typing data from 346 typists to analyze the effect of letter

₃₅ uncertainty on response times. In essence, every letter typed represents one choice reaction

₃₆ test trial from Hick's experiment. In Hick's experiment, the probability of each light turning

₃₇ on was manipulated to change the H value. This is analagous to the probability of certain

₃₈ letters appearing at a specific position in a word. Peter Norvig analyzed words from

₃₉ numerous texts and has determined the probability for each letter appearing at a specific

40 position in 2 to 9 letter words.

41   For example, the first letter of a two letter long word is has a 9.44% chance to be an

42 "a", 14.9% chance to be a "t", and a 25% chance to be an "i". The third letter of a five letter

43 long word is has a 13.5% chance as an "e", 4.91% chance as a "t", and a 11.4 % chance as a

44 "a". The H indices for each letter position - word length pair can be calculated. The H for

45 the 1st Letter, 2-letter-long word (abbreviated 1:2) is 2.85. The H for the 3rd Letter,

46 5-letter-long word (abbreviated 3:5) is 3.94. The differences in H tell us that the first letter

47 of a two letter long word has lower uncertainty ie. it is more predictable than the third letter

48 of a five letter long word.

49   This combination of letter probabilities creates an H value for each unique pair of letter

50 position and word length.

51   One potential problem that can hinder our experiment's ability to test the Hick's Law

52 is the fact that our typists vary in terms of their skill level. This means that differences in

53 RT between subjects will generate noise that may hide the H effect on RTs. An advantage of

54 testing Hick's Law with this data set is that the stimulus response coding is non-arbitrary.

55 The letters that typists see on the screen correspond exactly to the letters on their keyboards.

Methods

56

57   The uncertainty for each letter position in 1-9 letter long words was calculated using

58 Norvig's ngrams1.xlsx from his website:http://norvig.com/mayzner.html.and using the

59 equation H = N sigma (PlnP) where N is the number of alternatives (26 possible letters) and

60 P is probability of a given letter occurring at a specific letter position in a word of specific

61 length.

62   As an example, I demonstrate calculation of the H value of the 2nd letter of 3 letter

63 words: H = N sigma (PlnP) Total number of 3 letter words observed by Norvig: 1.52* 10^11

64 three letter words P("a")2 position = $3.1 10^{10}$ *three letter words with "a" as 2nd letter/1.52*

65 10^11 three letter words P("b")2 position = $3.64 10^{9}$ *three letter words with "b" as 2nd*

66 *letter/1.52* 10ˆ11 three letter words : : P("z")2 position $= 3.00 10^7$ *three letter words with*

67 *"z" as 2nd letter/1.52* 10ˆ11 three letter words H = 26 possible alternatives * sigma (as

68 letter from "a" to "z") P"letter"2position *ln P"letter"2position H = 2.85 bits

69 The relationship between uncertainty and letter position is visualized in Graph #. We

70 observe a trend of increasing uncertainty near the middle letter position of most words.

Data analysis and pre-processing

72 We used R (Version 3.4.2; R Core Team, 2017) and the R-packages *bindrcpp* (Version

73 0.2.2; Müller, 2018), *Crump* (Version 1.0; Crump, 2017), *data.table* (Version 1.10.4.3; Dowle

74 & Srinivasan, 2017), *dplyr* (Version 0.7.4; Wickham, Francois, Henry, & Müller, 2017),

75 *ggplot2* (Version 2.2.1; Wickham, 2009), *knitr* (Version 1.20; Xie, 2015), *papaja* (Version

76 0.1.0.9655; Aust & Barth, 2018), *Rcpp* (Eddelbuettel, 2017; Eddelbuettel & Balamuta, 2017;

77 Version 0.12.16; Eddelbuettel & François, 2011), *RcppZiggurat* (Version 0.1.4; Eddelbuettel,

78 2017), and *Rfast* (Version 1.8.8; Papadakis et al., 2018) for all our analyses.

79 For each subject, we recorded timestamps for each keystroke using JavaScript. We

80 applied the following pre-processing steps. We included IKSIs only for keystrokes involving a

81 lower case letter, and only for correct keystrokes that were preceded by a correct keystroke.

82 Outlier IKSIs were removed for each subject, on a cell-by-cell basis, using Van Selst &

83 Jolicoeur's (1994) non-recursive moving criterion procedure, which eliminated approximately

84 X% of IKSIs from further analysis.

Results

86 For each subject, we calculated mean IKSIs as a function of letter position and word

87 length. The letter position and word length factors were not factorially crossed. To

88 determine whether there were differences among the means we submitted the means to a

89 single factor repeated measures design with 45 levels (e.g., letter position|word length: 1|1,

90 1|2, 2|2, . . . 9|9). Figure 1 shows mean IKSIs collapsed over subjects, as a function of letter

91 position and word length.

92   The omnibus test indicated differences among the means were significantly different

93   from chance, F (44,15180) = 276.74, MSE = 1,269.66, p < .001. Visual inspection of figure

94   shows several trends across the means consistent with first-letter slowing and mid-word

95   slowing reported by Ostry (1983).

96   Our more important aim was to determine whether variation among these means can

97   be explained by variation in letter uncertainty. For this reason we do not exhaustively

98   discuss all of the possible 990 differences among these 45 conditions. Nevertheless, we did

99   conduct all 990 comparisons using bonferroni corrected paired samples t-tests. The results

100   are displayed Figure 2, which shows absolute mean differences between conditions color

101   coded for significance (light grey is significant).

Discussion

103   The takeaway from this experiment was that a typist types a letter faster when the

104   letters commonly found at the given letter position are few. These findings are consistent

105   with Hick's Law. As the alternatives become equiprobable, as it is observed in the middle of

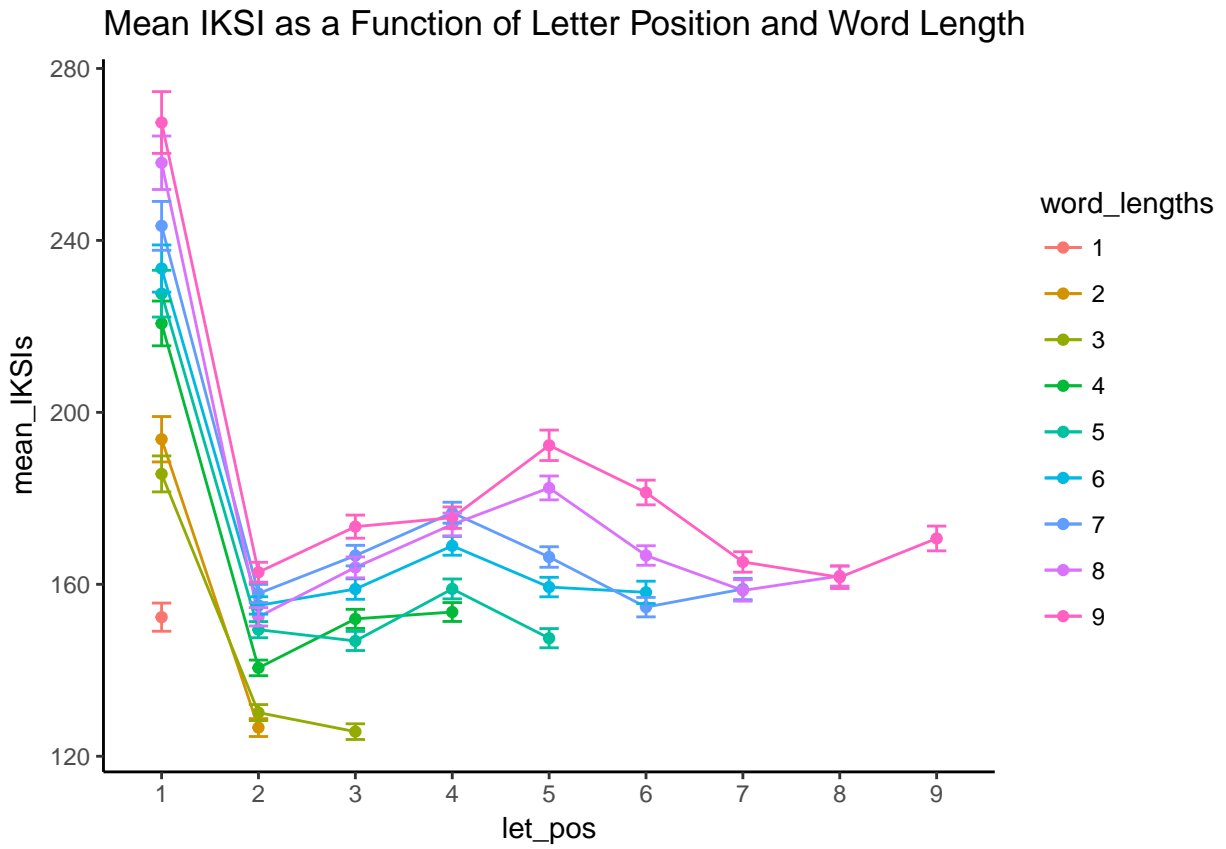106   words, the number of bits required to process the choices increases.

Conclusion

108   Future studies should investigate the role of probability of repetition in regulating

109   response times.

110   Kornblum in 1969 found that with constant H, trials that had higher chances of

111   sequentially repeating stimuli experienced faster response times.

<div style="text-align:center">**References**</div>

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown.*
    Retrieved from https://github.com/crsh/papaja

Crump, M. (2017). *Crump: Crump lab functions.*

Dowle, M., & Srinivasan, A. (2017). *Data.table: Extension of 'data.frame'.* Retrieved from
    https://CRAN.R-project.org/package=data.table

Eddelbuettel, D. (2017). *RcppZiggurat: 'Rcpp' integration of different "ziggurat" normal rng*
    implementations. Retrieved from
    https://CRAN.R-project.org/package=RcppZiggurat

Eddelbuettel, D., & Balamuta, J. J. (2017). Extending extitR with extitC++: A Brief
    Introduction to extitRcpp. *PeerJ Preprints*, *5*, e3188v1.
    doi:10.7287/peerj.preprints.3188v1

Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of*
    Statistical Software, *40*(8), 1–18. doi:10.18637/jss.v040.i08

Müller, K. (2018). *Bindrcpp: An 'rcpp' interface to active bindings.* Retrieved from
    https://CRAN.R-project.org/package=bindrcpp

Papadakis, M., Tsagris, M., Dimitriadis, M., Fafalios, S., Tsamardinos, I., Fasiolo, M., . . .
    Lakiotaki, K. (2018). *Rfast: A collection of efficient and extremely fast r functions.*
    Retrieved from https://CRAN.R-project.org/package=Rfast

R Core Team. (2017). *R: A language and environment for statistical computing.* Vienna,
    Austria: R Foundation for Statistical Computing. Retrieved from
    https://www.R-project.org/

Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York.
    Retrieved from http://ggplot2.org

Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). *Dplyr: A grammar of data*
    manipulation. Retrieved from https://CRAN.R-project.org/package=dplyr

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, Florida:

139   Chapman; Hall/CRC. Retrieved from https://yihui.name/knitr/

Mean IKSI as a Function of Letter Position and Word Length

Figure 1

Mean Absolute Differences

Figure 2