1 Ambiguous Words

2 Nicholas R. Harp[1], Catherine C. Brown[1], & Maital Neta[1]

3 [1] University of Nebraska-Lincoln

4 Author Note

5 Nicholas R. Harp, Department of Psychology, Center for Brain, Biology, and Behavior,

6 University of Nebraska-Lincoln Catherine C. Brown, Department of Psychology, Center for

7 Brain, Biology, and Behavior, University of Nebraska-Lincoln Maital Neta, Department of

8 Psychology, Center for Brain, Biology, and Behavior, University of Nebraska-Lincoln

9 Correspondence concerning this article should be addressed to Nicholas R. Harp,

10 Postal address. E-mail: nharp@huskers.unl.edu

Abstract

We found some ambiugous words.

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: ambiguity

Word count: X

21   Ambiguous Words

Introduction

23   We wanted to identify ambiguous words. Mention in this section why we also
24 generated clearly positive and negative words.

Study 1: Pilot

Methods

27   **Participants.**   Workers on Amazon's Mchanical Turk (MTurk) were invted to
28 participate in an eligibility screener worth \$0.20 with the option to earn a bonus of \$2.05 if
29 they met the requirements and completed the entire study. See **Supplementary**
30 **Information** for specific MTurk batch settings. The Workers clicked a hyperlink that
31 directed them to the study. The screener task included demographic questions and one block
32 of word ratings that included 5 instances of the word "negative" and 5 instances of the word
33 "positive" (see Procedure below for full details). Workers were invited to complete the entire
34 study if they indicated that they were over 18 years old, had English as their native
35 language, had no history of psychological or neurological disorder, and correctly rated the
36 words "positive" and "negative" as positive or negative with at least 80% accuracy. Of the
37 145 Workers who completed the screener, 119 met the eligibility requirements, and 103
38 (54.37% female, 45.63% male) chose to complete the entire study. The final sample was
39 3.88% Asian, 5.83% Black, 85.44% White, with a mean(sd) age of 37.16(10.60).

40   **Material.**

41   *Stimuli.*

42   We compiled an initial set of 59 words that we believed had two distinct definitions,
43 one clearly positive definition and one clearly negative definition. To create lists of clearly
44 positive and clearly negative words, we first created a master list of words that were included
45 in both the study by Warriner, Kuperman, and Brysbaert (2013), for valence and arousal

ratings, and the Enlgish Lexicon Project online word query (Balota et al., 2007), for lexical characterisic measurements. We then elimiated any words with a mean arousal rating that was greater than 1 standard deviation away from the mean arousal of the list of 59 ambiguous words. We classified "positive" words as those with a mean valence > 7 on the 1-9 scale used by Warriner et al. (2013); "negative" words had mean valence < 3. To ensure that all words shared similar lexical characteristics, we eliminated any words from the master list whose lexical characteristics did not fall within the minimum and maximum values of the 59 ambiguous words' lexical characteristics. The following were used for the cutoffs: length, the frequency of a word as reported by the Hyperspace Analogue to Language (HAL) study (Lund & Burgess, 1996), the log of HAL frequency, number of phonemes, number of syllables, number of morphemes, lexical decision reaction time and accuracy, and naming reaction time and accuracy. The final list of pilot words included 59 ambiguous, 267 positive, and 304 negative words.

All of the calculations described in this section were scripted using R version 3.6.1 and are available in the **Supplementary Information**.

Software.

All tasks were created and presented using Gorilla Experiment Builder (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2019). The study was only accessible to participants using a computer (not a phone or tablet) within the United States.

Procedure.

Screener and word rating task.

After giving informed consent, participants first answered demographic questions about their gender, age, race, native language, and whether they had ever been diagnosed with a psychological or neurological disorder. They then were shown a brief self-guided instructional walkthrough of the task before completing the screener.

71   Using a random seed, we selected 20 positive and 20 negative words from the final pilot

72   list for use in the screener task. These 40 words, along with 5 instances of the word

73   "positive" and 5 instances of the word "negative" were presented randomly, one at a time,

74   each following a 250 ms fixation cross. Each word remained on screen until the participant

75   indicated that they thought it was positive or negative by pressing A or L on their keyboard

76   (key pairing randomized across participants). If no response was made after 2000ms, a

77   reminder appeared on screen, "Please respond as quickly as you can! A = POSITIVE. L =

78   NEGATIVE." Participants who rated the words "positive" and "negative" with less than

79   80% accuracy were compensated for their time but were not invited to complete the rest of

80   the study. Participants were also excluded at this point if they indicated that they were

81   younger than 18, that English was not their native language, or that they had been

82   diagnosed with a psychological or neurological disorder.

83   The remaining 590 words from the final pilot list were randomly presented across 10

84   blocks of 59 words using the same button-press procedure as the screener block.

85   **Data preprocessing.**   Trials with a response time faster than 250ms (n =) were

86   removed from the data prior to analysis, as well as trials greater than 3 SDs above the mean

87   reaction time averaged across all trials.

88 **Results**

89   **Subjective ratings.**

90   **Reaction times.**

91   **Study 2: Comparison of words with valence bias and IPANAT**

92 **Methods**

93   **Participants.**

94   **Material.**

95   *Stimuli.*

96   *Valence Bias with Words.*

97   *Valence Bias with Faces.*

98   *Valence Bias with IAPS.*

99   *IPANAT.*

100   ***Software.***

101   **Procedure.**

102   ***Valence Bias Tasks.***

103   ***IPANAT.***

104   **Data analysis.**

105   ***Valence Bias Tasks.***

106   ***IPANAT.***

107   **Results**

108   **Subjective ratings.**

109   *Valence Bias with Words.*

110   *Valence Bias with Faces.*

111   *Valence Bias with IAPS.*

112   *IPANAT.*

113   **Reaction times.**

114   *Valence Bias with Words.*

115   *Valence Bias with Faces.*

116   *Valence Bias with IAPS.*

117   *IPANAT.*

118   **Relationships between the measures.**

119   **Discussion**

120   We did this study.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2019). Gorilla in our midst: An online behavioral experiment builder. *Behav Res.* https://doi.org/10.3758/s13428-019-01237-x

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*(3), 445–459. https://doi.org/10.3758/BF03193014

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*(2), 203–208. https://doi.org/10.3758/BF03204766

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav Res, 45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

₁₃₄ <div align="center">**Supplementary Information**</div>

₁₃₅ **MTurk Project Settings**

₁₃₆   The following are the settings used for the first batch on MTurk. This batch only

₁₃₇ contributed 6 respondents because the batch was published before the Gorilla task was fully

₁₃₈ functioning, and the batch expired before all HITs could be filled.

₁₃₉   • Title: Screener: Rate words as positive or negative (WARNING: This HIT may

₁₄₀   contain adult content. Worker discretion is advised.)

₁₄₁   • Description: Bonus available ($2.05) to those who meet eligibility. Complete short

₁₄₂   demographic questions. Use your keyboard to indicate if you think individual words

₁₄₃   are positive or negative.

₁₄₄   • Keywords: survey, demographics, rating, rate, words

₁₄₅   • Reward per response: $0.2

₁₄₆   • Number of respondents: 9

₁₄₇   • Time allotted per worker: 1 Hour

₁₄₈   • Survey expires in: 7 Days

₁₄₉   • Auto-approve and pay Workers in: 3 Days

₁₅₀   • Require that Workers be Masters to do your tasks: Yes

₁₅₁   • Specify any additional qualifications Workers must meet to work on your tasks:

₁₅₂   – Location is UNITED STATES (US)

₁₅₃   – HIT Approval Rate (%) for all Requesters' HITs greater than 95

₁₅₄   – Number of HITs Approved greater than 5000

₁₅₅   • Project contains adult content: selected

₁₅₆   • Task Visibility: Hidden - Only Workers that meet my Qualification requirements can

₁₅₇   see and preview my tasks

₁₅₈   The same settings were used for the rest of the batches except that they did not

₁₅₉ require that Workers be Masters and the Number of HITs Approved was set to greater than

160   500, not 5000.

Study 1 Stimuli

162   Insert link to repository for "pick_words.R"