**Harrisburg University of Science and Technology**

**Institutional Review Board**

**APPLICATION TO USE HUMAN SUBJECTS IN RESEARCH**

**For Exempt Studies Only**

***IRB ETHICS TRAINING****: The participating Faculty member(s), Graduate Student Researcher(s), and any External Researcher(s) MUST have a valid completion certificate from the CITI Course in Human Subjects Online Training before submitting an IRB application for projects involving human subjects. Include a copy of your CITI Training Completion Certificate with your IRB application if one is not already on file.*

1. PROPOSED DATA COLLECTION DATES: From (09/01/2021) to (09/01/2022)

Data collection dates should allow time for the IRB to review your protocol. Please allow at least one (1) week from the date you turn in the application for processing.

2. INVESTIGATOR(S):Copy and paste additional investigator names as needed. *If an undergraduate student project, the faculty advisor should be listed as a Co-Investigator and as the approving faculty advisor).*

Investigator Name: Erin M. Buchanan

Program: ANLY

Email: [ebuchanan@harrisburgu.edu](mailto:ebuchanan@harrisburgu.edu)

Faculty Advisor Name: Click or tap here to enter text.

Program: Click or tap here to enter text.

Email: Click or tap here to enter text.

For all students, this research is for (*check all that apply*):

☐ Master’s Thesis/Project ☐ Independent Study

☐ PhD Dissertation  ☐ GRAD695 Course requirement

☐ Undergraduate Project ☐ Other: (describe project here)

3. **PROJECT TITLE**: Semantic Priming Across Many Languages (SPAML)

4. PARTICIPANTS:

a. Number of participants proposed/anticipated: 50,000

b. Type(s) of participants:

☐ Children (17 or younger) X Adults (18 years of age or older)

☐ Patients in institutions X HU students (18 years of age or older)

☐ Prisoners X Faculty or external collaborators

☐ Pregnant women X Other:

This project is in collaboration with the Psychological Science Accelerator. Each team that contributes to data collection will obtain their own exemptions (i.e., some countries do not have IRB regulations), reliance on the HU IRB, or their own IRB review. Here we are indicating that we will have external collaborators who will select participants in many different ways (their classrooms, paid samples, social networks, etc.).

5. FUNDING:

Are you seeking funding for this project/research? ☐ No X Yes

We plan to submit to the Prolific Research Competition, as well as the National Science Foundation in their linguistics category. Both of these are under preparation.

Does the funding agency require IRB approval? ☐ No X Yes ☐ N/A

They can rely on this IRB approval for the project.

6. REVIEW CATEGORY: **Please mark all items that apply**.

*Note: Research with children or pregnant women often cannot be reviewed under expedited review. Please consult with the IRB Administrator to see if your protocol involving these subjects would require full board review.*

**Exempt** **Review** (based on the following categories):

☐ Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

X Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, **unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.**

☐ Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available **or** if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

Less likely types in the exempt category include the following:

☐ Research and demonstration projects which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine: (i) Public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs; (iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs.

☐ Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture.

7. **ADDITIONAL INFORMATION REQUIRED**:

You may insert the required information below in this section OR include it in your submission as separate documents:

* 1. Project Topic or Research Question(s):

Semantic priming has been studied for nearly fifty years across various experimental manipulations and theoretical frameworks. Critically, the understanding of semantic priming relies on reliable, well-studied stimuli with defined similarity values. In the last twenty years, the publication rates of normed stimuli databases and corpora (i.e., large bodies of text) has exponentially increased. Further, newer computational models of concept representation have been detailed using these databases. Using these newer models, we can define similarity between concepts to create reliable stimuli for study in semantic priming. This research project will meet the need for a database of semantic priming values, particularly in non-English languages that are cross linguistically complete. Large-scale data in this area is sparse, unlike the other published databases found in Buchanan et al. [(2019a)](https://paperpile.com/c/d09zVa/NVFy/?noauthor=1). Therefore, this study aims to provide data that complements and extends the published data, which would encourage research on methodology, item characteristics, models, cross-language consistency in priming, and other theoretical areas that semantic priming has been applied to previously. The global aims of this project include:

1) Create an online framework to collect semantic priming data, modeled after the success of the Small World of Words project [(De Deyne et al., 2019)](https://paperpile.com/c/d09zVa/Sp2f). The online framework would allow data collection from any internet capable computer, thus lowering the burden on research labs to collect data of this nature. The online framework can then be used to deliver updates to the data, even after the conclusion of the initial data collection.

2) Provide a large dataset of response latencies and priming scores for prime and target words in up to 55 languages, as available on the Open Subtitles Project. Further, these prime and target words will be supplemented with variables that are theoretically important for research in cognitive architectures to provide a dataset with less missing data. The dataset provided allows researchers to continue to use these datasets to select carefully controlled stimuli, as well as investigate questions about items, participants, reliability, and language.

* 1. Explanation of why the data collected is anonymous.
     1. Methodology (the design of the study)
     2. Data Collection (what elements and how you will collect data)

**Method**

**Participants**

Data from the English Lexicon Project [(Balota et al., 2007)](https://paperpile.com/c/d09zVa/o46g) and the Semantic Priming Project [(Hutchison et al., 2013)](https://paperpile.com/c/d09zVa/mfwo) were used to estimate the minimum sample size necessary for the study. The aim of this study is to provide a large dataset, rather than test a hypothesis, so traditional ways to estimate sample size via power and effect size were not applicable. Therefore, an accuracy in parameter estimation approach was employed using the previous data as a metric. In this approach, one focuses on finding a confidence interval around a parameter that would be “sufficiently narrow” [(Kelley, 2007; Kelley, Darku, & Chattopadhyay, 2018; Maxwell, Kelley, & Rausch, 2008)](https://paperpile.com/c/d09zVa/J8iv+vG6z+ZIrT)**.** Both the English Lexicon Project and Semantic Priming Project used a lexical decision task, which will be employed in this study. These data were used to estimate the likely standard errors of lexical decision data for individual words. These values were used as the rubric of accurately measured lexical decision response latencies.

Given proposed standard error value, the data was then sampled with replacement to determine the sample size that would provide that standard error value. One hundred words within the data were selected, and samples starting at *n* = 5 to *n* = 400 were selected (increasing in units of five). The standard error for each of these samples was then calculated for the simulation, and the percent of samples with standard errors at or less than the estimated population value was then tabulated. From this calculation, *n* for each target concept was estimated at 100-320 participants. The design of the study, the number of words per session, expected data loss due to incorrect answers, number of target words desired, and number of required conditions were all taken into account and the final estimate for sample size per language is 741 to 4741. The complete code and description of this process is detailed at: [here](https://github.com/doomlab/SPAML/tree/master/parameter_estimation).

This sample size estimation represents a major improvement from previous database collection studies, as many have used the traditional *n* >= 30 as a way to guess at minimum sample size. As indicated, it’s often unclear how to exactly estimate a sample size for these types of studies, and this study will detail that procedure to provide guidance for future work. The upper range of estimated participants is high because of the uncertainty in estimating an “accurate” parameter. Because the variability of the sample size is quite large, we will employ a stopping procedure to ensure participant time and effort is maximized, and data collection is minimized. The minimum sample size will be 50 participants per concept or 741 total participants, and the maximum will be 320. After 50 participants, each concept will be examined for standard error, and data collection for that concept will be stopped when the standard error reaches an average of the two metrics found in this exploration (0.06, 0.012; see supplemental material) or 0.09. This process will be automated online and checked in a daily subroutine. From the current simulations, this approximates to 100-150 participants per word, and 1482-2223 participants per language total.

**Materials**

Semantic priming focuses on word-pair relatedness or similarity, and therefore, prime-target pairs are often chosen for their similarity in the related condition. The unrelated condition pairs are then created by shuffling the prime-target pairs so that the prime word is combined with a target word it has no relationship to. Non-words are created by changing one to two letters in a prime or target word to create a nonsense word (*nurse* → *lurse*), with the stipulation that they must be pronounceable and not pseudo-homophones (i.e., wherein the pronunciation sounds like a real word, *keep* → *keap*). Consequently, the choice of related words is key for the study. There are multiple measures of semantic similarity including the cosine between overlapping features [(Buchanan et al., 2019b)](https://paperpile.com/c/d09zVa/eXKH), free association probabilities [(De Deyne, Navarro, & Storms, 2013)](https://paperpile.com/c/d09zVa/C1nC), and local/global coherence values from network models [(Siew & Vitevitch, 2016; Vitevitch et al., 2014)](https://paperpile.com/c/d09zVa/jKpT+6ExH). However, the underlying data for these calculations is spotty across languages. Therefore, one solution is to use the Open Subtitles and subs2vec projects to calculate lexical co-occurrence as a measure of semantic similarity (Lison & Tiedemann, 2016; van Paridon & Thompson, 2020). With the subtitle data, we will take the first 10000 most frequent nouns, adjectives, adverbs, and verbs from each language, and these will be cross-referenced using the *translateR* package [(Lucas & Tingley, 2014)](https://paperpile.com/c/d09zVa/5RQM). Next, a distributional space model for each language will be used from the subs2vec project to identify concepts related to the 10000 most frequent words and to calculate their respective similarity values [(Mandera et al., 2017)](https://paperpile.com/c/d09zVa/hhqi). The top five most related words will be selected, and these will be cross-referenced across languages. Native speakers will be recruited to ensure the accurate translation of word pairs. The related word pairs (*n* = 1000) will be selected from the list using each concept only once, favoring pairs with translations in most languages. If a selected pair does not exist in a language, translation from a Native speaker will be used to create that pair. Words will also be cross referenced for polysemy (i.e., multiple meanings) and these will be restricted when possible. Lastly, concepts will be examined for their relative statistics on lexical measures (length, part of speech, neighborhood, phonemes/morphemes) and subjective measures (age of acquisition, imageability, concreteness, valence, dominance, arousal, and familiarity) because of their known associations with concept representation. Psuedowords will be created by replacing a random letter in the selected words while ensuring this letter matches potential bigrams found in the language. The code for this selection procedure and current wordlist can be found: [here](https://github.com/SemanticPriming/SPAML/tree/master/stimuli).

**Procedure**

A small demonstration of the experiment can be found at: [here](https://open-lab.online/code/PSA%20LDT%20Example/?generate=true). The study will be programmed using lab.js [(Henninger, Shevchenko, Mertens, Kieslich, & Hilbig, 2019)](https://paperpile.com/c/d09zVa/6aux), which is an online, open-source study creation project. Precise timing measurement is required for this study, and the lab.js team has documented the accuracy of measurement within their framework [(Henninger, Shevchenko, Mertens, Kieslich, & Hilbig, 2018)](https://paperpile.com/c/d09zVa/3a8Q), and previous work has shown no differences between lab and web-based data collection for response latencies [(Hilbig, 2016)](https://paperpile.com/c/d09zVa/1857). In addition, SPALEX, a large lexical decision database in Spanish was collected completely online [(Aguasvivas et al., 2018)](https://paperpile.com/c/d09zVa/JEZ6). We will recommend that research labs use Chrome as their browser, however, meta-information about the browser and operating system are saved when participants take the experiment to control for implementation differences. Participants will be directed to an online web portal to take the study, and all data will be retained in the online platform with nightly backups to GitHub. They will be asked to indicate their gender (male, female, other, prefer not to say), year of birth for age, and education level (none, elementary school, high school, bachelors, masters, doctorate) for demographic variables. To continue in the study, they will select their primary language, which will direct them to the appropriate stimuli set. The research lab ID will be collected (to track external collaborators) but no identifying information will be collected about participants (i.e., IP address or other information than listed here for demographics).

Participants will be required to complete the study on a computer, rather than a mobile or tablet device. This requirement allows for tracking of the display of the device which will indicate important aspects about screen size, browser, and timing accuracy. In order to enforce this requirement, participants will be asked to hit the spacebar to continue the study. Instructions on how to complete a lexical decision task will be shown on the next screen, followed by 10 practice trials. Each trial starts with a fixation cross (+) in the middle of the screen for 500 ms. The concept will then be displayed in the middle of the screen in uppercase San-Serif font (i.e., NURSE). On the bottom of the screen the answer choices will be shown as the traditional keys next to the *shift* key depending on the common keyboard layout for that language (i.e., Z and / on a QWERTY keyboard or > and - on a QWERTZ keyboard). These choices will be reversed in half of the subjects, which will be randomly selected at the start of the study to counterbalance word/nonword selection. Participants will enter their choice for each concept, and then the next word will appear with an intertrial interval of 500 ms (i.e., the time between the offset of the first concept and onset of the next concept, when the fixation cross is showing). Responses will time out after 5 seconds and move on to the next trial. After ten trials, participants will see the instruction screen again with a reminder that they will now be doing the real task.

After 100 trials, the participants will be shown a short break screen with the option to continue by hitting the spacebar after 10 seconds. After six blocks of 100 trials (600 words), the experiment will end with a thank you screen. On this screen, participants will indicate what type of credit they are receiving for the study (course credit, payment), and they will be given instructions on how to indicate they have completed the study to the appropriate lab. Participants will be allowed to take the study multiple times (see below). These values will be customized based on data collection type (i.e., Mechanical Turk, participant pool, etc.). An estimate for the amount of time required for the study is approximately twenty to thirty minutes including practice trials, instructions, and breaks. We will pilot test the number of stimuli to keep the study under 30 minutes and will lower the number of trials accordingly.

A primary goal of this project is to provide a complete dataset of priming and other important related linguistic variables. Lexical measures, such as length, frequency, part of speech, and the number of phonemes (i.e., sounds in a word) are easily created from the concept or the SUBTLEX projects. Subjective measures are concept characteristics that are rated by participants, such as age of acquisition (approximate age you learned a concept), imageability (how easy the concept is to imagine), concreteness (how concrete is the concept), valence (emotion), arousal, dominance (controlled versus dominated), and familiarity. For concepts that are missing these values in a target language, participants will be asked to provide ratings on a single metric (i.e., they would only see instructions for familiarity or arousal). Each participant will be asked to provide 25-50 ratings of concepts, given the need for a particular language, while also controlling for the length of the task to prevent fatigue in the experiment. These will only be presented at the end of the experiment to prevent interactions with priming effects. We will use the available large databases of these variables to estimate sample size necessary for these ratings using the same simulation procedure detailed above.

8. **AFFIRMATION OF COMPLIANCE:**

**Investigators or researchers are required to notify the IRB of substantive changes to protocol, unanticipated adverse, serious events experienced by participants, and project completion. Projects lasting longer than one year require an annual Request for Continuation (Protocol Renewal) or Notice of Project Completion. Failure to submit may result in adverse actions IAW IRB Policy. All consent forms and data must be kept at least three years after the study ends.**

***I agree to follow the procedures outlined herein and to ensure that the rights and welfare of human participants are properly protected. I will commence the study only after receiving approval from the IRB) and having complied with required modifications. I will promptly report additions, changes, or problems involving the rights or welfare of human participants to the IRB by contacting the IRB Chair. If the project continues for more than one year from the approval date, I will submit the required documentation.***

***I affirm that I have read and reviewed the accuracy of this application and accept responsibility for the ethical conduct of this research, supervision of human participants, and maintenance of data and informed consent documentation as required by the IRB.***

erin m. buchanan ebuchanan@harrisburgu.edu

Signature of Investigator HU E-mail Address Date (06/15/2021)

*(Written or digital)*

Click or tap here to enter text.

Signature of Co-investigator HU E-mail Address Date (MM/DD/YYYY)

*(Written or digital)*

*(Cut and Paste additional investigator signature lines as needed).*

**APPROVAL OF FACULTY ADVISOR OR SPONSOR (if a student project or thesis):**

***I affirm that I have read and reviewed the accuracy of this application and accept responsibility for the ethical conduct of this research, supervision of human participants, and maintenance of data and informed consent documentation as required by the IRB.***

***I agree to follow the procedures outlined herein for my student(s) and to ensure that the rights and welfare of human participants are properly protected. I will ensure the study does not commence until the study has been approved by the HU IRB and have complied with required modifications. I will promptly report additions, changes, or problems involving the rights or welfare of human participants to the IRB by contacting the IRB Chair. If the project continues for more than one year from the approval date, I will submit the required documentation.***

*Copy and paste additional faculty advisor approval signatures and contact information lines as needed below.)*

Printed Name of Faculty Advisor: Click or tap here to enter text.

Program: Click or tap here to enter text.

Phone: Click or tap here to enter text.

HU E-mail Address: Click or tap here to enter text.

Signature of Faculty Advisor Date (MM/DD/YYYY)

*(Written or digital)*

CONSENT FORM

Understanding Word Processing and Meaning

You are invited to be in a research study about how you read and process words, along with their meaning. We ask that you read this form and ask any questions you may have before agreeing to be in the study.

This study is being conducted by Dr. Erin M. Buchanan, Professor of Cognitive Analytics at Harrisburg University of Science and Technology.

### Background Information:

In this study, you will be asked to complete different questions about word concepts. For example, you may be asked to define a word’s characteristics, rate how familiar you are with a word, or simply judge if a string of letters is a real word.

**Procedures:**

You will take this study entirely online from a desktop or laptop computer with a keyboard. You will be given instructions about the experiment sections which are randomly selected for each person. After you complete the experiment, you can learn more about the study and goals of the research. The entire study should take less than thirty minutes to complete.

### Risks and Benefits of being in the Study:

No identifying information will be collected from you, and therefore, your responses should be anonymous. The current study is similar to an online game, which may cause some fatigue or boredom based on the task you are asked to complete.

There is no direct benefit to you for participating in this study. However, your responses will contribute to our understanding of language and cognitive memory processes.

**Compensation:**

You may be compensated when taking part in this study through your local researcher.

**Confidentiality and Data Sharing:**

Measures are taken to ensure that all information you provide will be anonymous. The data from this project will be posted publicly for other researchers to use; however, no data will be directly linked to you. Your name or other identifying information will not be entered into the dataset and no references will be made in verbal or written reports that could link you to the study. In any publication, information will be provided in such a way that you cannot be identified.

Before your data are shared outside the research team, any potentially identifying information will be removed. The anonymous data may be used by the research team or shared with other researchers, for both related and unrelated research purposes in the future. Your anonymous data may also be made available in online data repositories such as the Open Science Framework (which are free data repositories that require registration to have access), which allow other researchers and interested parties to access the data for further analysis.

**Voluntary Nature of the Study:**

**Participation in this study is voluntary:**

Your decision whether or not to participate will not affect your current or future relations with Harrisburg University of Science and Technology or your local institution. If you decide to participate, you are free to not answer any question or withdraw at any time without affecting those relationships.

**Contacts and Questions:**

The researchers conducting this study are Dr. Erin M. Buchanan in partnership with the Psychological Science Accelerator. You may ask any questions you have now. If you have questions later, **you are encouraged** to contact Dr. Erin M. Buchanan at [ebuchanan@harrisburgu.edu](mailto:ebuchanan@harrisburgu.edu).

**Questions or Concerns:**

This study has been reviewed by Harrisburg University of Science and Technology’s Institutional Review Board (IRB). The IRB has determined that this study fulfills the human research subject protections obligations required by state and federal law and University policies.

***A copy of this information to keep for your records will be provided upon request.***

**Study instructions:**

**Demographics:**

What is your native (first) language? (this choice will direct them to the full study in their native language).

Please tell us a little bit about yourself.

Please tell us your gender: male female other prefer not to say

Which year were you born? Please enter a four-digit year:

Please tell us your education level: less than High School diploma, High School, Associates or two year degree, University or four year degree, Masters Degree, Doctoral degree (please note these will be made culturally relevant for each language)

**Semantic priming task:**

This experiment is concerned with how people process words. You will be asked to view words and judge them. The words will be presented in the middle of the screen. You should ask yourself if the word on the screen is a real word or a fake word. If you recognize the word on the screen like **COLD**, please press the **KEY** key for real word. If the word is made-up nonsense, like **WERM**, please press the **KEY** key for fake word. During practice, we will give you feedback. We would like to ask you to respond as quickly as you can while remaining accurate. You have five seconds to respond to every word. Please press the **SPACE BAR** to try a few for practice.

(note the **KEY** will be manipulated between participants and based on the traditional keyboard for that language placing it by the Z and ? keys on a traditional QWERTY keyboard).

Please press **KEY** for a real word, and **KEY** for a fake word. (shown while competing the priming task wherein words are shown on the screen)

Great job! Remember you should use the **KEY** key for real words, and the **KEY** key for fake words. You will now complete the real task. Remember to go as quickly as possible while getting the correct answer. You will not see any feedback during this section. Please press the **SPACE BAR** to continue. (shown after practice and before the real trials).

Please take a short break on this screen. You may press the **SPACE BAR** when you are ready to continue. (shown after each 50-100 pairs to combat fatigue).

Thank you for completing the first part of the experiment. On the next screen, you will judge words for some of their properties to help us measure how these words are understood by native speakers. (a random section below will be shown).

**Word meaning task: (McRae et al., 2005)**

We want to know how people read words for meaning. Please fill in features of the word that you can think of. Examples of different types of features would be: how it looks, sounds, smells, feels, or tastes; what it is made of; what it is used for; and where it comes from. Here is an example:

duck: is a bird, is an animal, waddles, flies, migrates, lays eggs, quacks, swims, has wings, has a beak, has webbed feet, has feathers, lives in ponds, lives in water, hunted by people, is edible

Complete this questionnaire reasonably quickly, but try to list at least a few properties for each word. Thank you very much for completing this questionnaire.

**Age of acquisition task: (Kuperman et al., 2012)**

Please rate the following word by entering the age in years at which you think you learned the word. We mean the age at which you would have understood the word if somebody had used it in front of you, EVEN IF YOU DID NOT use, read or write it at the time. If you do not know the word, please enter the **X** key.

**Concreteness task: (Brysbaert et al., 2014)**

Some words refer to things or actions in reality, which you can experience directly through one of the five senses. We call these words concrete words. Other words refer to meanings that cannot be experienced directly but which we know because the meanings can be defined by other words. These are abstract words. Still other words fall in-between the two extremes, because we can experience them to some extent and in addition we rely on language to understand them. We want you to indicate how concrete the meaning of each word is for you by using a 5-point rating scale going from abstract to concrete. A concrete word comes with a higher rating and refers to something that exists in reality; you can have immediate experience of it through your senses (smelling, tasting, touching, hearing, seeing) and the actions you do. The easiest way to explain a word is by pointing to it or by demonstrating it (e.g. To explain 'sweet' you could have someone eat sugar; To explain 'jump' you could simply jump up and down or show people a movie clip about someone jumping up and down; To explain 'couch', you could point to a couch or show a picture of a couch). An abstract word comes with a lower rating and refers to something you cannot experience directly through your senses or actions. Its meaning depends on language. The easiest way to explain it is by using other words (e.g. There is no simple way to demonstrate 'justice'; but we can explain the meaning of the word by using other words that capture parts of its meaning). Because we are collecting values for all the words in a dictionary (over 60 thousand in total), you will see that there are various types of words, even single letters. Always think of how concrete (experience based) the meaning of the word is to you. In all likelihood, you will encounter several words you do not know well enough to give a useful rating. This is informative to us too, as in our research we only want to use words known to people. We may also include one or two fake words which cannot be known by you. Please indicate when you don't know a word by using the letter N (or n).

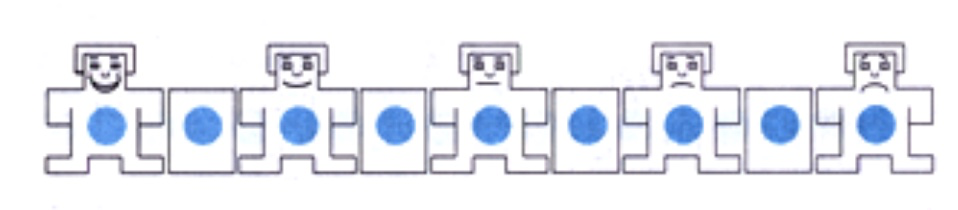
So, we ask you to use a 5-point rating scale going from abstract to concrete and to use the letter N when you do not know the word well enough to give an answer.

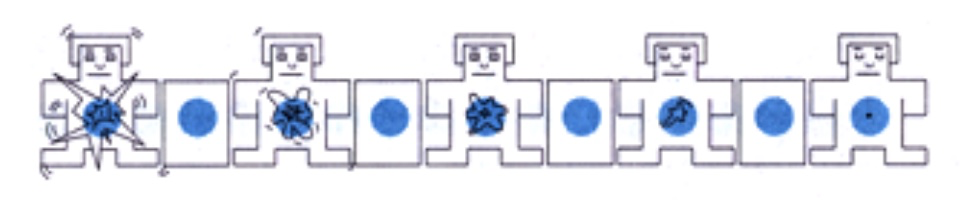
Abstract (language based) Concrete (experience based)

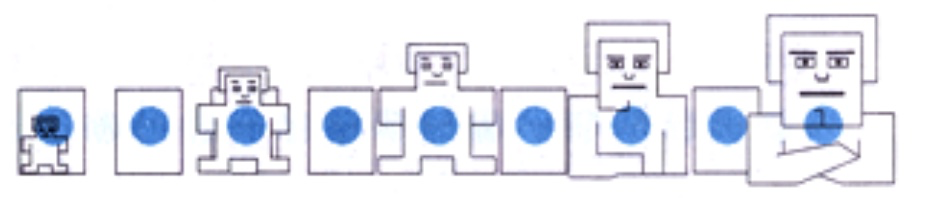
1 2 3 4 5

N = I do not know this word well enough to give a rating.

**Valence, Arousal, and Dominance task: (Bradley & Lang, 1999)**

**valence**

**arousal**

**dominance**

The study being conducted today is investigating emotion, and how people respond to different types of words.

We call this set of figures SAM, and you will be using these figures to rate how you felt while reading each word. SAM shows three different kinds of feelings: Happy vs. Unhappy, Excited vs. Calm, and Controlled vs. In-control. You will use these scales to make all 3 ratings for each word that you read. Please notice that each of the three feelings are arrayed along a different scale. The left panel shows the happy-unhappy scale, which ranges from a smile to a frown. At one extreme of this scale, you are happy, pleased, satisfied, contented, hopeful. When you feel completely happy you should indicate this by selecting the 1 on the left. The other end of the scale is when you feel completely unhappy, annoyed,unsatisfied, melancholic, despaired, or bored. You can indicate feeling completely unhappy by selecting the 9 on the right. The figures also allow you to describe intermediate feelings of pleasure, by selecting any values in the middle. If you feel completely neutral, neither happy nor sad, select the 4 in the middle.

The excited or calm scale is the second type of feeling displayed here. At one extreme of this scale you are stimulated, excited, frenzied, jittery, wide-awake, or aroused.When you feel completely aroused, select the 1 on the left. Now look at the other end of the excited-calm scale, which is the completely opposite feeling. Here you would feel completely relaxed,calm, sluggish, dull, sleepy, or unaroused. Indicate feeling calm by selecting the 9 on the right. As with the happy-unhappy scale, you can represent intermediate levels of excitedness or calmness by selecting any other number. If you are not excited nor at all calm, select the 4 in the middle.

The last scale of feeling which you will rate is whether you felt controlled or in control. At one end of the scale you have feelings characterized as completely controlled, influenced, cared-for, awed, submissive, orguided. Please indicate feeling controlled by selecting the 1 on the left. At the opposite end of this scale, you would select the 9 if you feel completely in control, influential, important, dominant, autonomous, or controlling. You can indicate feeling dominant selecting the 9 on the right. Note that when the figure is large, you feel in control, and that it will be very small when you feel controlled. If you feel neither in control nor controlled you should select the 4.

Please work at a rapid place and don’t spend too much time thinking about each word. Rather, make your ratings based on your first and immediate reaction as you read each word.

**Familiarity task:**

We are interested in determining your familiarity with a large set of words. We will present different words on the following pages. Your task is to rate your familiarity with each word on a 7-point rating scale by circling a number from 1 to 7.

Use this guide to help you in deciding which number to circle:

7 - You recognize the word and are confident that you know the meaning of the word.

6 - You think you know the meaning of the word but are not certain that the meaning you know is correct.

5 - You are certain that you have seen the word but you only have a vague idea of its meaning.

4 - You recognize the word as one you have seen or heard before, but you don't know the meaning of the word.

3 - You are pretty sure that you have seen or heard the word but you are not positive.

2 - You think that you might have seen or heard the word before.

1 - You have never seen or heard the word before.

We don't expect you to know or even recognize all of these words. Don't be concerned if there are many words you have never seen or heard before. These words reflect a very wide sample of words in LANGUAGE and it is expected that many of the words will be unknown to most people. Please be sure to respond to every word.

**Iconicity task:**

For this task, we want to know how iconic you think words are. Some words are considered to be iconic; the word somehow sounds like what it means. For example, haha sounds like the sound made when you laugh. You might be able to guess its meaning even if you did not know the language. Other words are not iconic at all; for example, there is nothing ‘occasional’ or ‘frequent’ sounding about the word. If you did not know the language, you would not be able to guess their meanings. For each word that you will see, rate on a scale of 1 to 7 how iconic you think the word is, with 1 as not iconic at all and 7 as highly iconic. Say the words aloud before making their rating. For example, haha is very iconic and would be a 7; tree not at all iconic and would be a 1. Words that are intermediate in iconicity, of course, should be rated appropriately between the two extremes, for example with 3 or 4. If you do not know the meaning of a word, put in X. Work fairly quickly but be as accurate as possible in your ratings. Feel free to use the entire range of numbers, from 1 to 7; at the same time, don’t be concerned about how often you use a particular number as long as you are honest in your ratings.