# NRC-CNRC

# **Norms of Valence, Arousal, and Dominance** for 20,000 English Words

- > Has fine-grained real-valued scores for V, A, and D (core dimensions of meaning)
- > We show that the annotations are reliable (high split-half reliability scores)
- > Has applications in NLP, psychology, digital humanities, etc.
- > We show that demographic attributes impact how we view the world

# Core Dimensions of Connotative Meaning (Osgood et al., 1957; Russell, 1980, 2003)

valence (V): positive/pleasure – negative/displeasure
arousal (A): active/stimulated – sluggish/bored
dominance (D): powerful/strong – powerless/weak



# Saif M. Mohammad

National Research Council Canada



🖂 saif.mohammad@nrc-cnrc.gc.ca 🛛 🕑 @SaifMMohammad

Areas of Interest:

Computational Linguistics, Machine Learning, Psycholinguistics Emotions, Creativity, and Fairness in Language

## **Our Lexicons**

#### The NRC Valence, Arousal, and Dominance Lexicon

- ratings of V, A, and D for ~20,000 words
- http://saifmohammad.com/WebPages/nrc-vad.html

Thus, we can compare words on the V, A, D dimensions:

- banquet indicates more positiveness than funeral
- *nervous* indicates more arousal than *lazy*
- queen indicates more dominance than delicate

#### Existing VAD Lexicons Created using Rating Scales

ANEW (Bradley and Lang, 1999), ~1,000 words Warriner et al. Norms (2013), ~14,000 words

fixed granularity; scale region bias

poor consistency (intra- and inter-annotators)



#### **The NRC Emotion Lexicon**

- associations for ~14,000 words with eight emotions
- http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

### The NRC Emotion Intensity Lexicon

- intensity scores for ~6000 words with anger, fear, joy, sadness
- http://saifmohammad.com/WebPages/AffectIntensity.htm

#### The NRC Word–Colour Association Lexicon

- associations for ~14,000 words with 11 common colours
- http://saifmohammad.com/WebPages/lexicons.html

# Comparative Annotations: Best–Worst Scaling (BWS) (Louviere & Woodworth, 1990)

- The annotator is presented with four words and asked:
   which word is associated with the most/highest X
  - (happiness / activeness / powerfulness)
- which word is associated with the least/lowest X
- Five out of the six inequalities are known:
  - if A is highest and D is lowest, then: A > B, A > C, A > D, B > D, C > D

# Annotation

- 20K common English terms by crowdsourcing (each tuple by six people)
- There separate questionnaires for V, A, D; ~800K annotations

### Entries with Highest and Lowest Scores. Range: 0 (lowest V/A/D) to 1 (highest V/A/D).

	Hig	Highest		Lowest	
Dimension	word	score		word	score
valence	love	1.000		toxic	0.008
	happy	1.000		nightmare	0.005
arousal	abduction	0.990		mellow	0.069
	exorcism	0.980		siesta	0.046
dominance	powerful	0.991		frail	0.069
	leadership	0.981		weak	0.045

Real-valued scores (Orme, 2009):
 score(w) = %best(w) - %worst(w) (transformed to a 0 to 1 range)

#### Advantages:

- preserves the comparative nature
- keeps the number of annotations down to about 2N
- leads to more reliable, less biased, more discriminating annotations (Kiritchenko and Mohammad, 2017)

# **Reliability (Reproducibility) of the VAD Annotations**

Average split-half reliability (SHR): approach to determine consistency (Kuder and Richardson, 1937; Cronbach, 1946)



Pearson correlation: -1(most inversely correlated) to 1(most correlated)

This study was approved by the NRC Research Ethics Board (NRC-REB) under protocol number 2017-98. REB review seeks to ensure that research projects involving humans as participants meet Canadian standards of ethics.



## Split-Half Reliability Scores for VAD Annotations

Annotations	# Terms	# Annotations	V	A	D
Warriner et al. (2013)	13,915	20 per term	0.91	0.79	0.77
Ours (Warriner terms)	13,915	6 per tuple	0.95	0.91	0.91
Ours (all terms)	20,007	6 per tuple	0.95	0.90	0.90

These SHR scores show that highly reliable fine-grained ratings can be obtained. Also:

- our V–D correlation is 0.48; Warriner V–D correlation is 0.71
- correlations of the scores across the two lexicons are low: V–V: 0.81; A–A: 0.62; D–D: 0.33

# 🙁 🚯 😰 🍣 🔕 🤤

## Shared Understanding of VAD Judgments by Demographic and Personality Trait Groups

#### Groups with significantly higher within-group agreement shown below

		Valence	Arousal	Dominance
Gender:	F-F vs. M-M	M-M	F—F	M-M
Age:	Y-Y vs. G-G	G—G	G—G	Y—Y
Big 5 Traits:	Ag-Ag vs. Di-Di	Ag—Ag	Ag—Ag	Di—Di
	Co–Co vs. Ea–Ea	-	Co-Co	Co-Co
	Ex-Ex vs. In-In	Ex-Ex	Ex-Ex	Ex-Ex
	Ne-Ne vs Se-Se	Se-Se	-	Se-Se
	Op-Op vs CI-CI	Op—Op	Op—Op	Op—Op

women: higher shared understanding of A of terms

- men and those that are secure: higher shared V and D
- over 35 years: higher shared V and A
- extroverts and open to experiences: higher shared V, A, D

= female	M = male	Ag = Agreeableness (friendly and compassionate)
′ = young (≤35)	G = grownups (>35)	Di = Disagreeableness (careful in whom to trust, argumentative)

Ne = Neurotic (often feel anger, anxiety, vulnerability) Co = Conscientiousness (efficient and organized) Se = Secure (rarely feel anger, anxiety, vulnerability) Ea = Easygoing (easy-going and carefree)

Op = Open to experiences (inventive and curious) Ex = Extrovert (outgoing, energetic, seek the company of others) CI = Closed to experiences (consistent and cautious) In = Introvert (solitary, reserved, meeting many people causes anxiety)



