

## Studying the Impact of Annotation Perspective and Representation Format on Dimensional Emotion Analysis



**Sven Buechel and Udo Hahn**

Jena University Language & Information Engineering (JULIE) Lab  
Friedrich-Schiller-Universität Jena, Jena, Germany  
{sven.buechel, udo.hahn}@uni-jena.de  
<http://www.julielab.de>



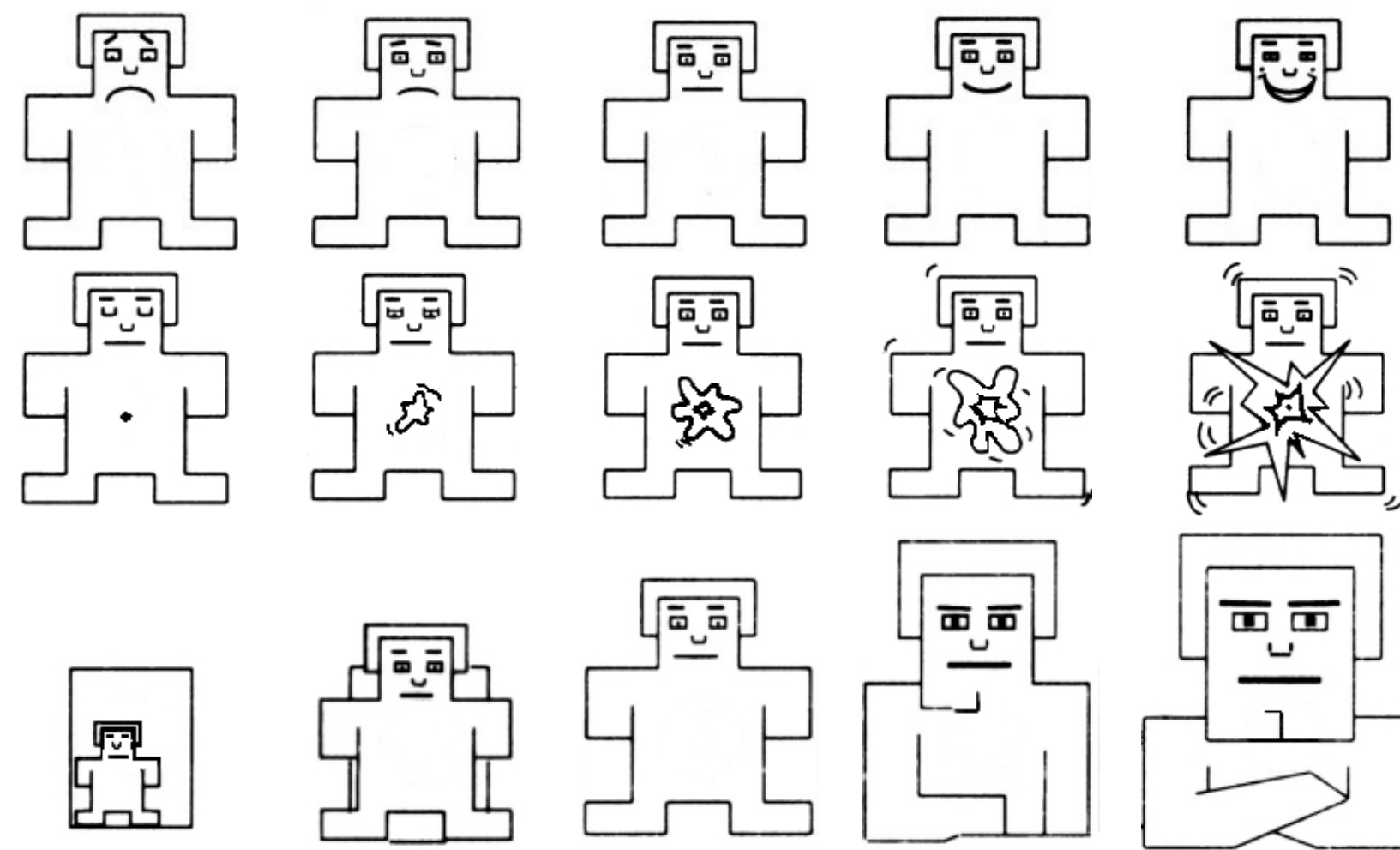
### Motivation

- Build large-scale gold standard for novel emotion representation (VAD)
- Compare reader and writer perspective
- Enable mapping between emotion formats

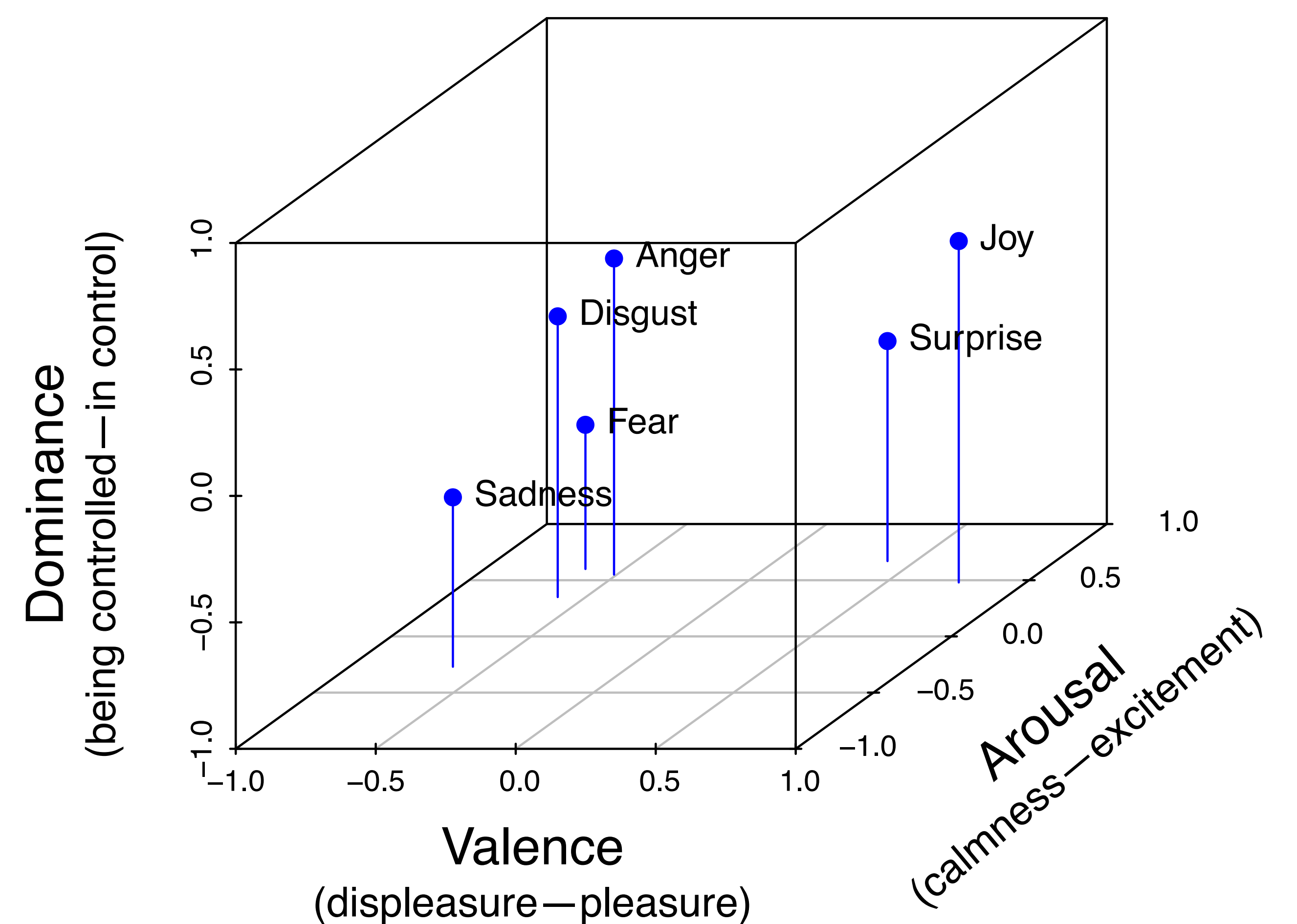
### Measuring Emotion

Self-Assessment Manikin

(Bradley & Lang, 1994)



### Valence-Arousal-Dominance (VAD)



### Corpus Acquisition

- Raw data:
  - MASC  $\subset$  ANC (Ide et al., 2010)
  - SemEval-2007 Task 14 (Strapparava & Mihalcea, 2007)
  - Genre-balanced
- Annotate full corpus according to reader and writer perspective. Pilot: Buechel & Hahn (2017)
- 5 raters per sentence and perspective (CrowdFlower)

Genre Distribution

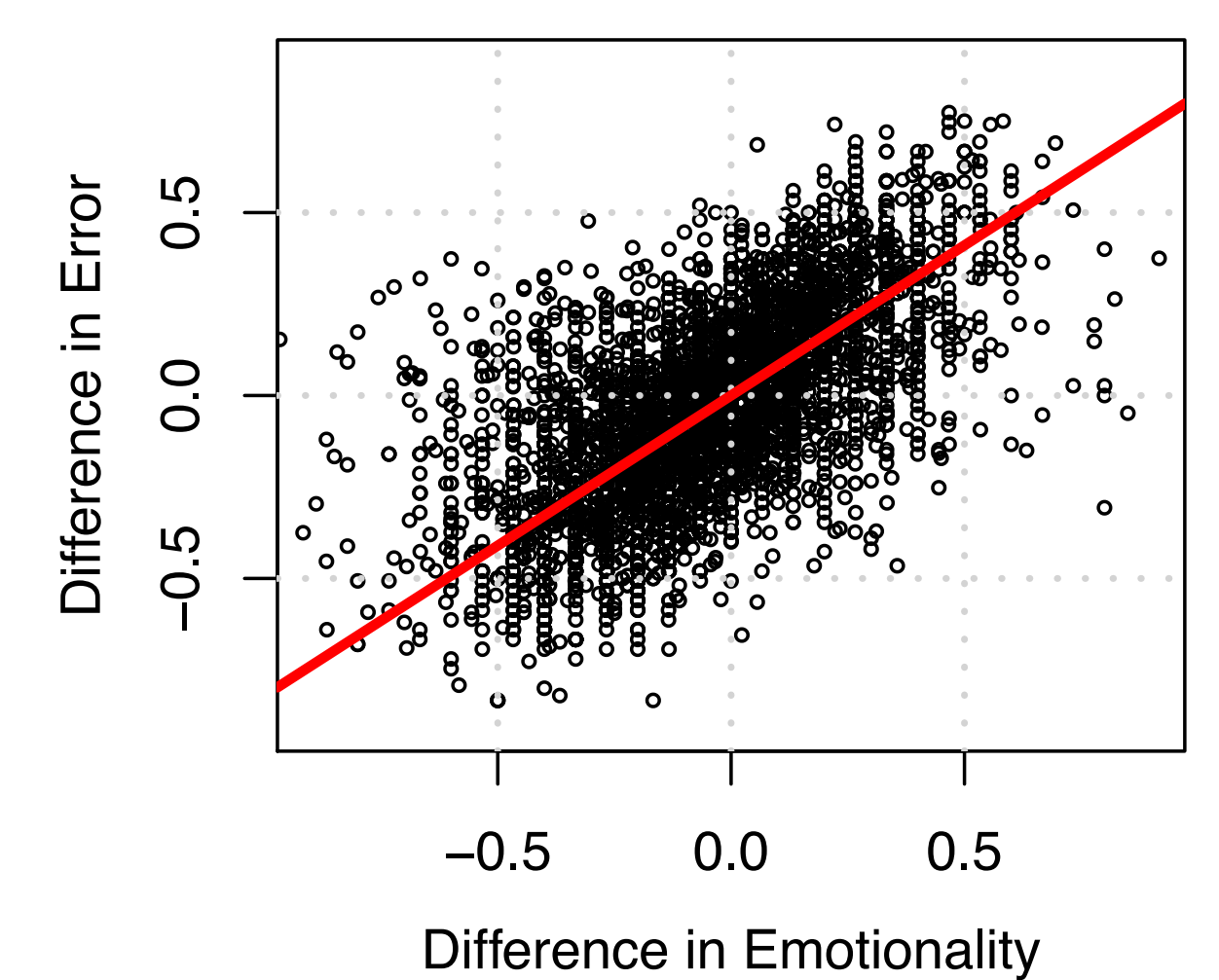
Corpus	Domain	Raw	Filtered
MASC	blogs	1,378	1,336
	essays	1,196	1,135
	fiction	2,893	2,753
	letters	1,479	1,413
	newspapers	1,381	1,314
	travel guides	971	919
SemEval07	news headlines	1,250	1,192
<b>Sum</b>		<b>10,548</b>	<b>10,062</b>

### Bi-Perspectival Design

Correlation- ( $r$ ) and Error- (MAE) based IAA

	Valence	Arousal	Dominance	Average
$r$ -writer	0.698	0.578	0.540	0.605
$r$ -reader	0.738	0.595	0.570	0.634
MAE-writer	0.300	0.388	0.316	0.335
MAE-reader	0.349	0.441	0.367	0.386

- Reader: better correlation but worse error IAA
- Also more emotional ratings
- Emotionality correlates with error
- Increased error explained by higher intensity



### Bi-Representational Design

- Previous studies hard to compare (incompatible formats)
- Can we automatically map between formats?
- Train kNN models to predict Basic Emotions given VAD
- Writer and reader combined reaches human IAA

	Joy	Anger	Sad.	Fear	Disg.	Surp.	Avg.
IAA	.60	.50	.68	.64	.45	.36	.54
Writer	.68	.40	.67	.47	.27	.15	.44
Reader	.73	.47	.68	.54	.36	.15	.49
Writer&Reader	.78	.50	.74	.56	.36	.17	.52
Diff Writer – IAA	+0.08	-.10	-.01	-.17	-.17	-.21	-.09
Diff Reader – IAA	+0.13	-.03	+0.00	-.10	-.09	-.22	-.05
Diff W&R – IAA	+0.18	+0.00	+0.05	-.08	-.09	-.19	-.02

### Conclusion

- Largest multi-annotated emotion corpus
  - for more than one perspective
  - for more than one emotion format
  - at all (10k sentences)
- Reader perspective turned out superior
- Mapping VAD to Basic Emotions reaches human annotation capacity
- Available: <https://github.com/JULIELab/EmoBank>

### References

- Bradley & Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59.
- Buechel & Hahn. 2017. Readers vs. writers vs. texts: Coping with different perspectives of text understanding in emotion annotation. In *LAW 2017*.
- Ide, Baker, Fellbaum & Passonneau. 2010. The Manually Annotated Sub-Corpus: A community resource for and by the people. In *ACL 2010*.
- Strapparava & Mihalcea. 2007. SemEval-2007 Task 14: Affective text. In *SemEval 2007*.