



Modeling Empathy and Distress in Reaction to News Stories

Sven Buechel ^{2*} Anneke Buffone ^{1*}

Barry Slaff ¹

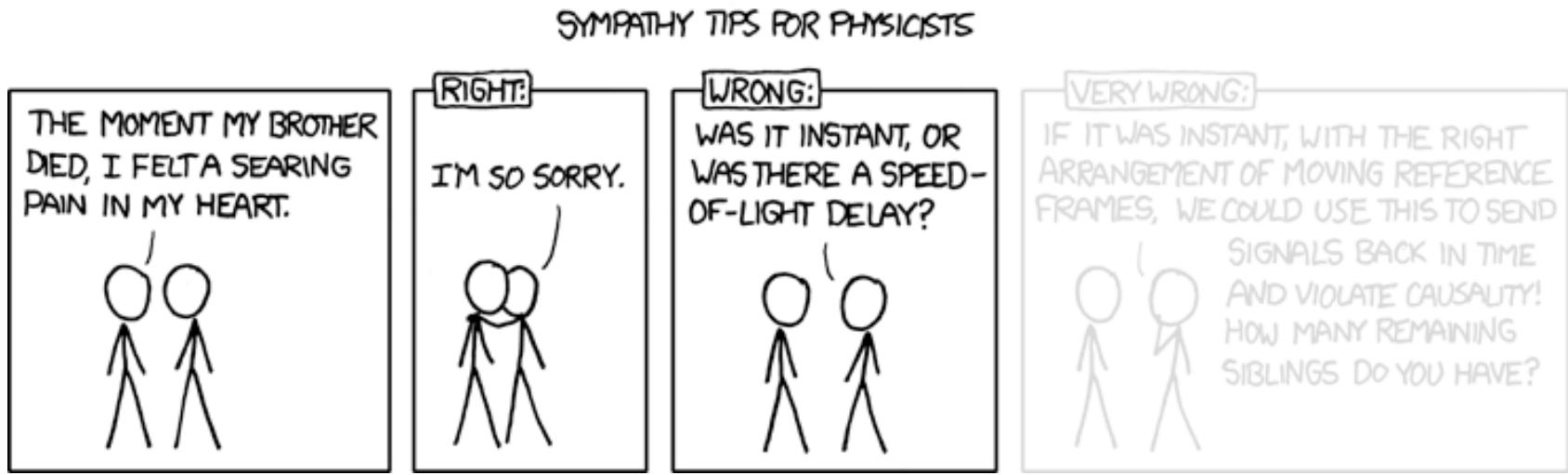
Lyle Ungar ¹

João Sedoc ¹

* equal contribution

¹ World Well-Being Project
University of Pennsylvania
<http://www.wwbp.org>

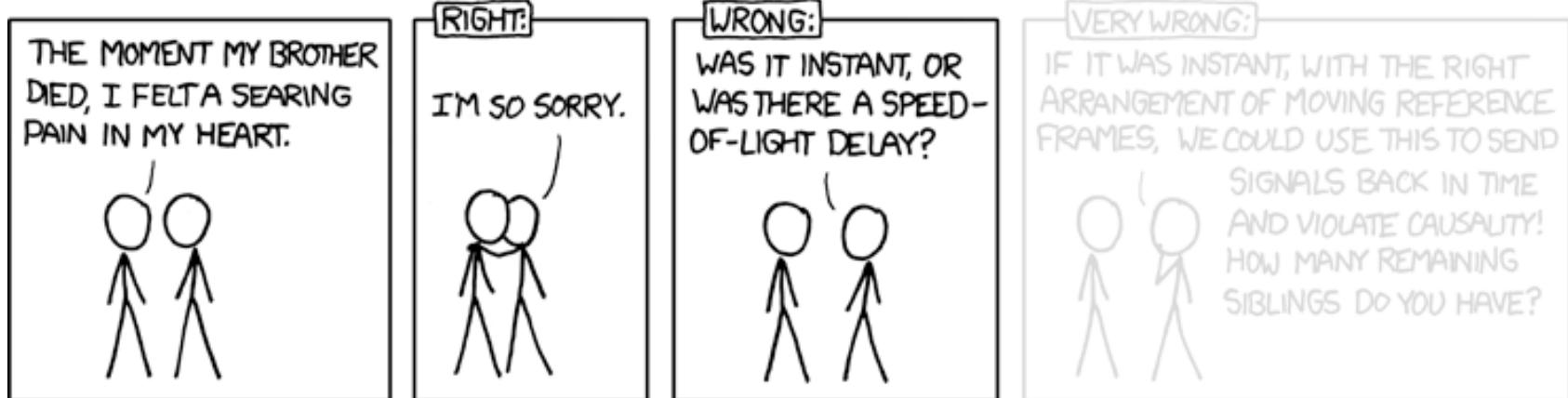
² JULIE Lab
Friedrich-Schiller-Universität Jena
<http://www.julielab.de>



<https://xkcd.com/660/> CC-BY-NC 2.5

EMPATHY TIPS FOR ENGINEERS

~~SYMPATHY TIPS FOR PHYSICISTS~~



<https://xkcd.com/660/> CC-BY-NC 2.5

Why Empathy?

- Crucial for how we experience the world and communicate within it
- Great potential for social media analysis

Why Empathy?

- Crucial for how we experience the world and communicate within it
- Great potential for social media analysis
- Scarce work on empathy modeling in written language

Shortcomings of Prior Work

- No publicly available gold standard
- 3rd person ground truth
- Disconnected from psychological theory and research

Contribution

- No publicly available gold standard
- 3rd person ground truth
- Disconnected from psychological theory and research

Contribution

- No publicly available gold standard
 - First publicly available gold standard (CC-BY)
- 3rd person ground truth
- Disconnected from psychological theory and research

Contribution

- No publicly available gold standard
 - First publicly available gold standard (CC-BY)
- 3rd person ground truth
 - Annotations by the experienter
(new annotation methodology)
- Disconnected from psychological theory and research

Contribution

- No publicly available gold standard
 - First publicly available gold standard (CC-BY)
- 3rd person ground truth
 - Annotations by the experienter
(new annotation methodology)
- Disconnected from psychological theory and research
 - Distinguish two different types of empathy
(in line with psych. research)

Empathic Concern vs. Personal Distress

- Consensus in psych. that there are multiple forms of empathy
- We follow most popular distinction by Batson et al. (1987)
 - **Empathic Concern**
positive, other-focused
 - **Personal Distress**
negative, self-focused

Empathic Concern vs. Personal Distress

- Consensus in psych. that there are multiple forms of empathy
- We follow most popular distinction by Batson et al. (1987)
 - **Empathic Concern**
positive, other-focused
 - **Personal Distress**
negative, self-focused
- Post-hoc analysis shows that both are distinct ($r=.45$)

Problem Definition

Message	Empathy	Distress
<i>I'm sorry to hear about Dakota's parents. [...]</i>	4.8	3.1

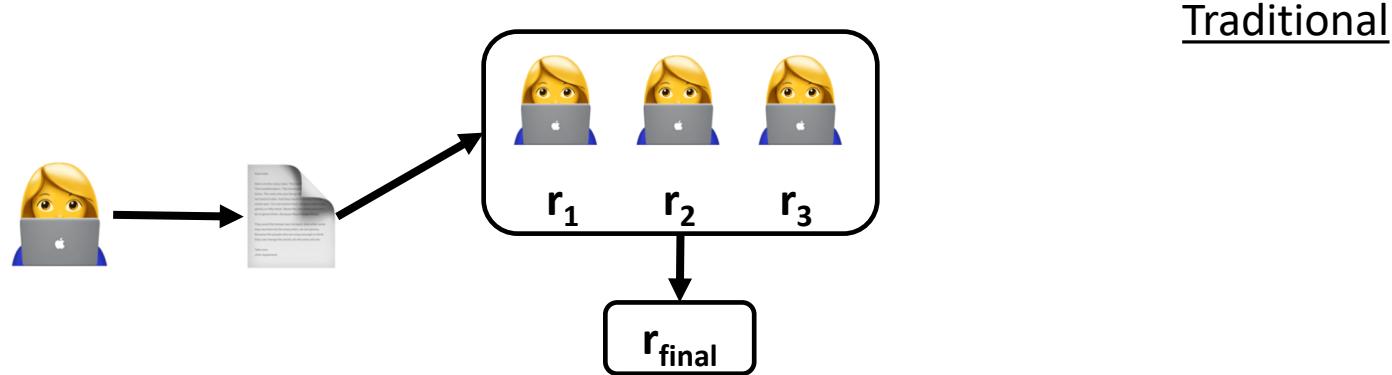
- Given a natural language utterance ...
- ... predict empathy and distress **of the writer** from [1, 7]

Annotation Methodology

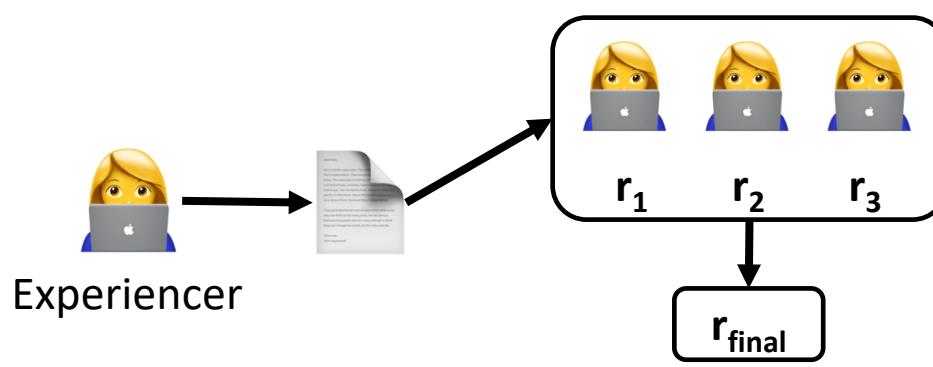
Traditional



Annotation Methodology



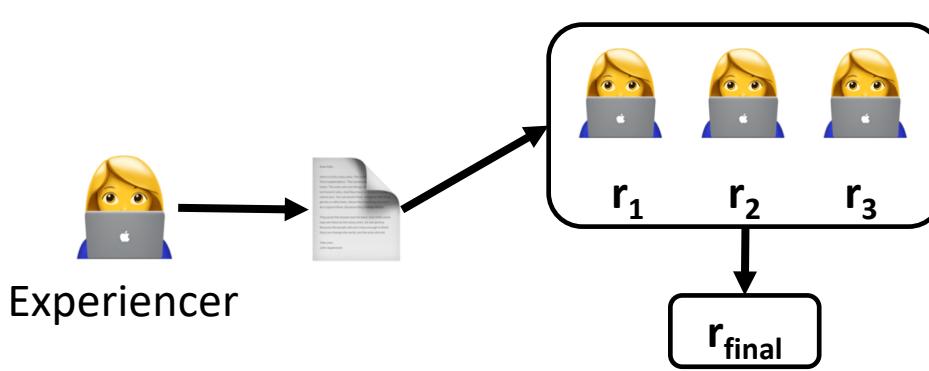
Annotation Methodology



Traditional

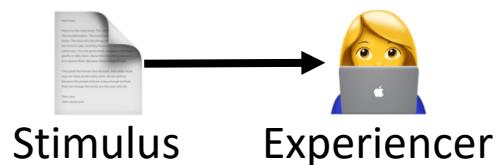
- Annotators “guess” experiencer’s feelings
- Models biased towards annotators

Annotation Methodology



Traditional

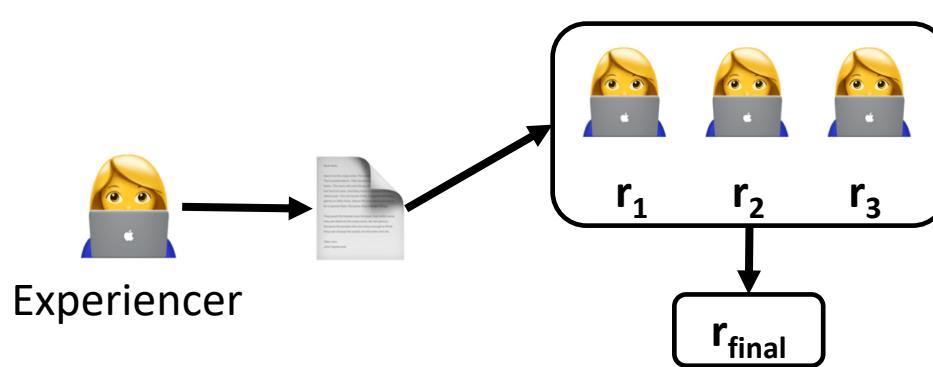
- Annotators “guess” experiencer’s feelings
- Models biased towards annotators



Proposed

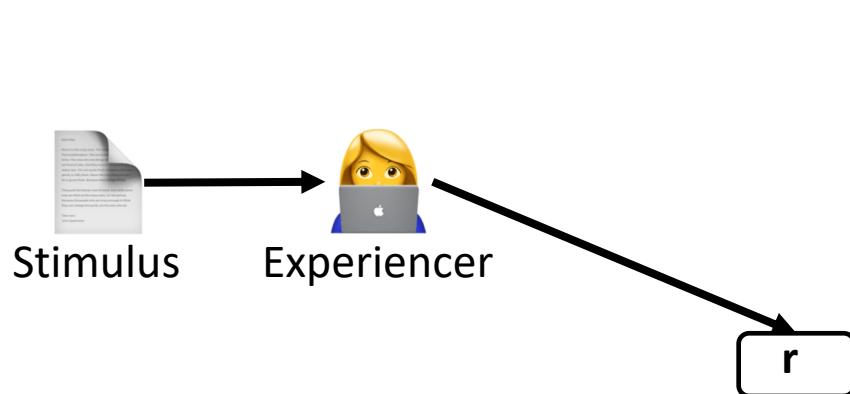
- Annotation as psych. experiment

Annotation Methodology



Traditional

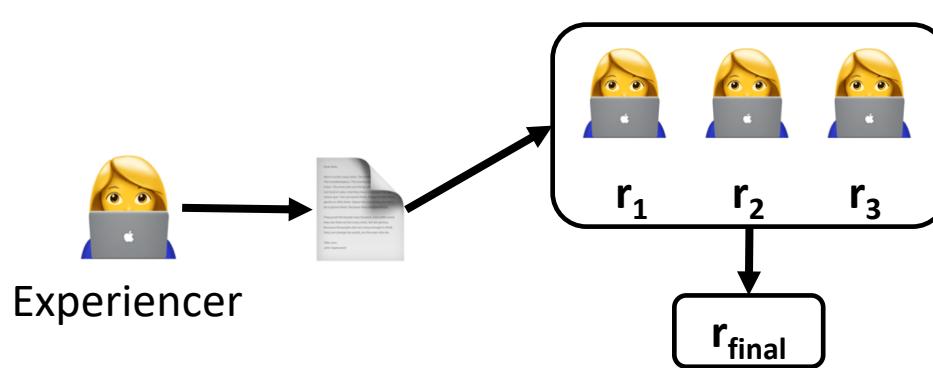
- Annotators “guess” experiencer’s feelings
- Models biased towards annotators



Proposed

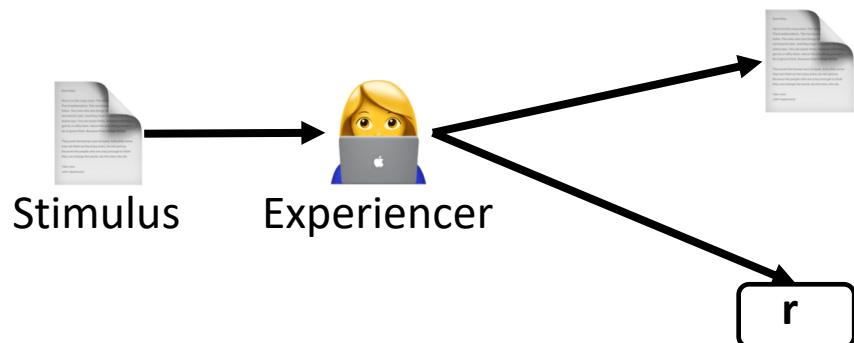
- Annotation as psych. experiment

Annotation Methodology



Traditional

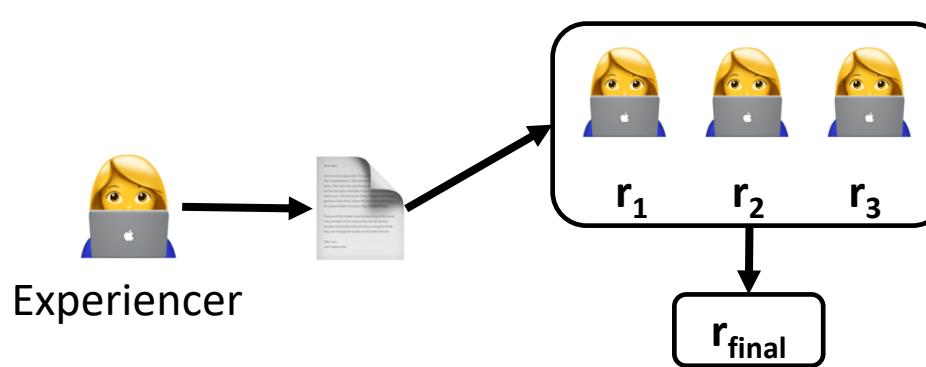
- Annotators “guess” experiencer’s feelings
- Models biased towards annotators



Proposed

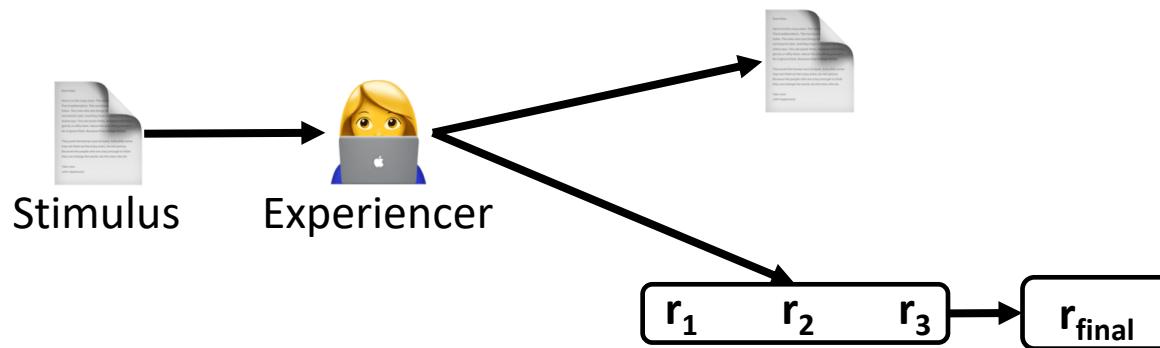
- Annotation as psych. experiment
- Experiencer produces **message and rating**

Annotation Methodology



Traditional

- Annotators “guess” experiencer’s feelings
- Models biased towards annotators



Proposed

- Annotation as psych. experiment
- Experiencer produces **message and rating**
- Reliable through **multi-item scales**

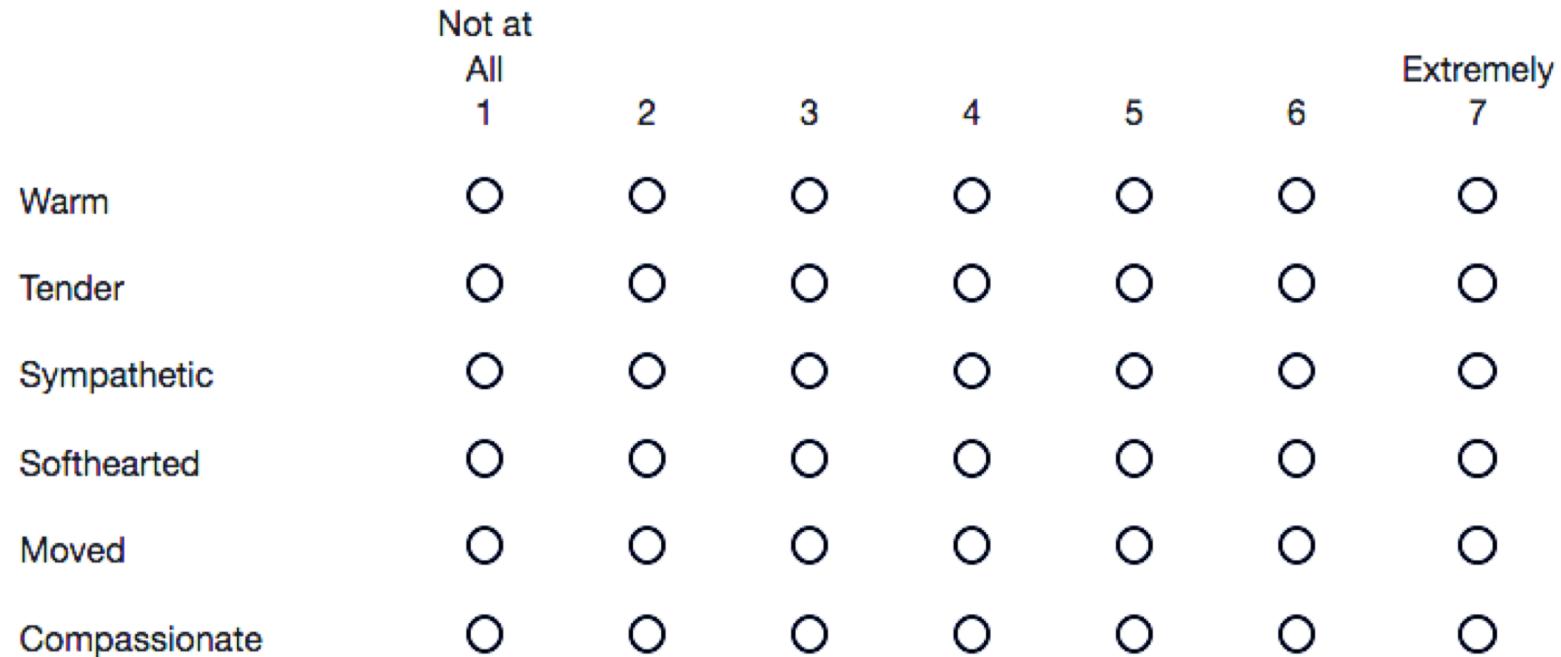
Use of Multi-Item Scales

How strongly do you feel the following emotion? Using the 1-7 scale below, please indicate your agreement.

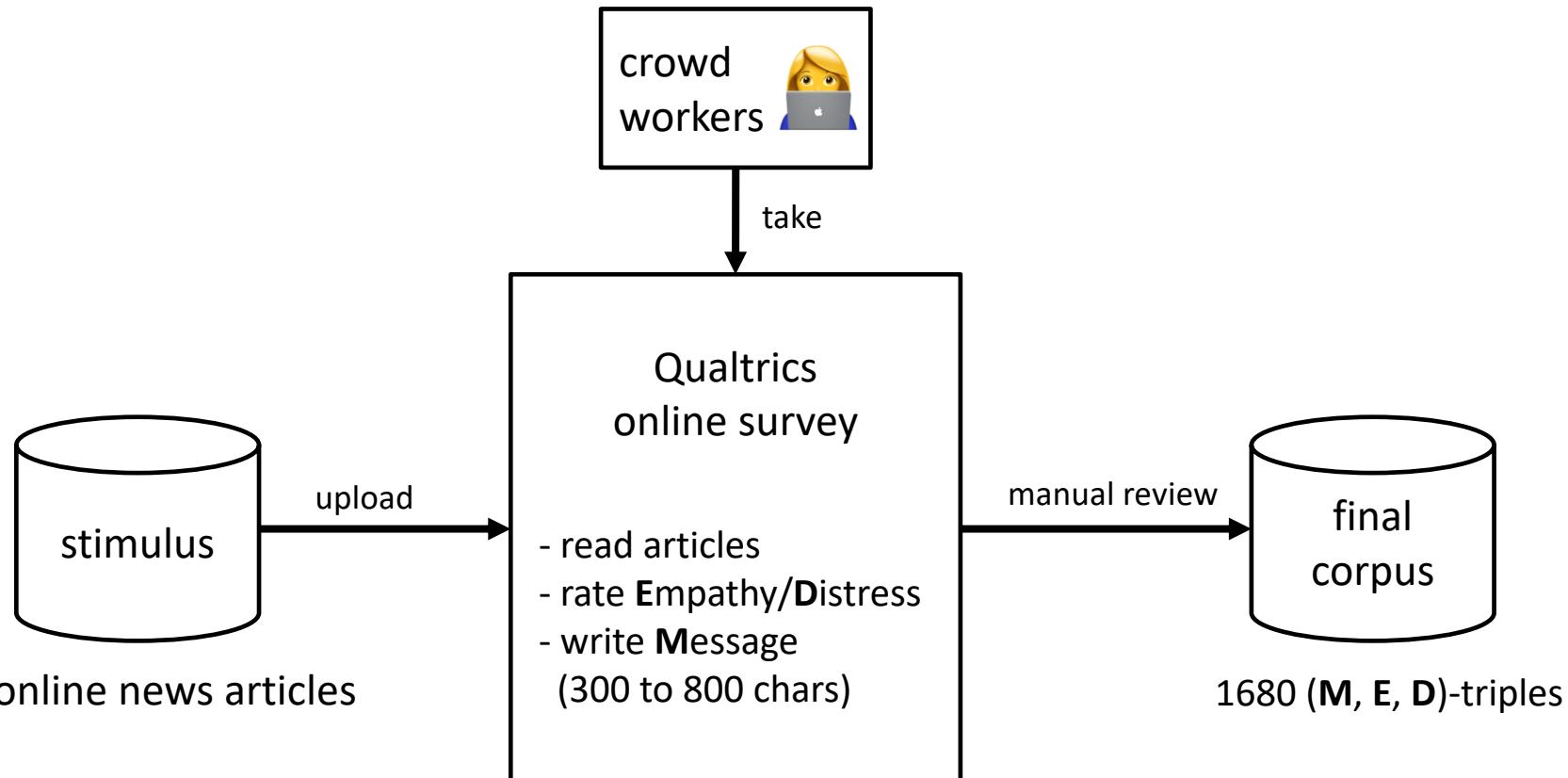
	Not at All	1	2	3	4	5	6	Extremely 7
Warm	<input type="radio"/>							
Tender	<input type="radio"/>							
Sympathetic	<input type="radio"/>							
Softhearted	<input type="radio"/>							
Moved	<input type="radio"/>							
Compassionate	<input type="radio"/>							

Use of Multi-Item Scales

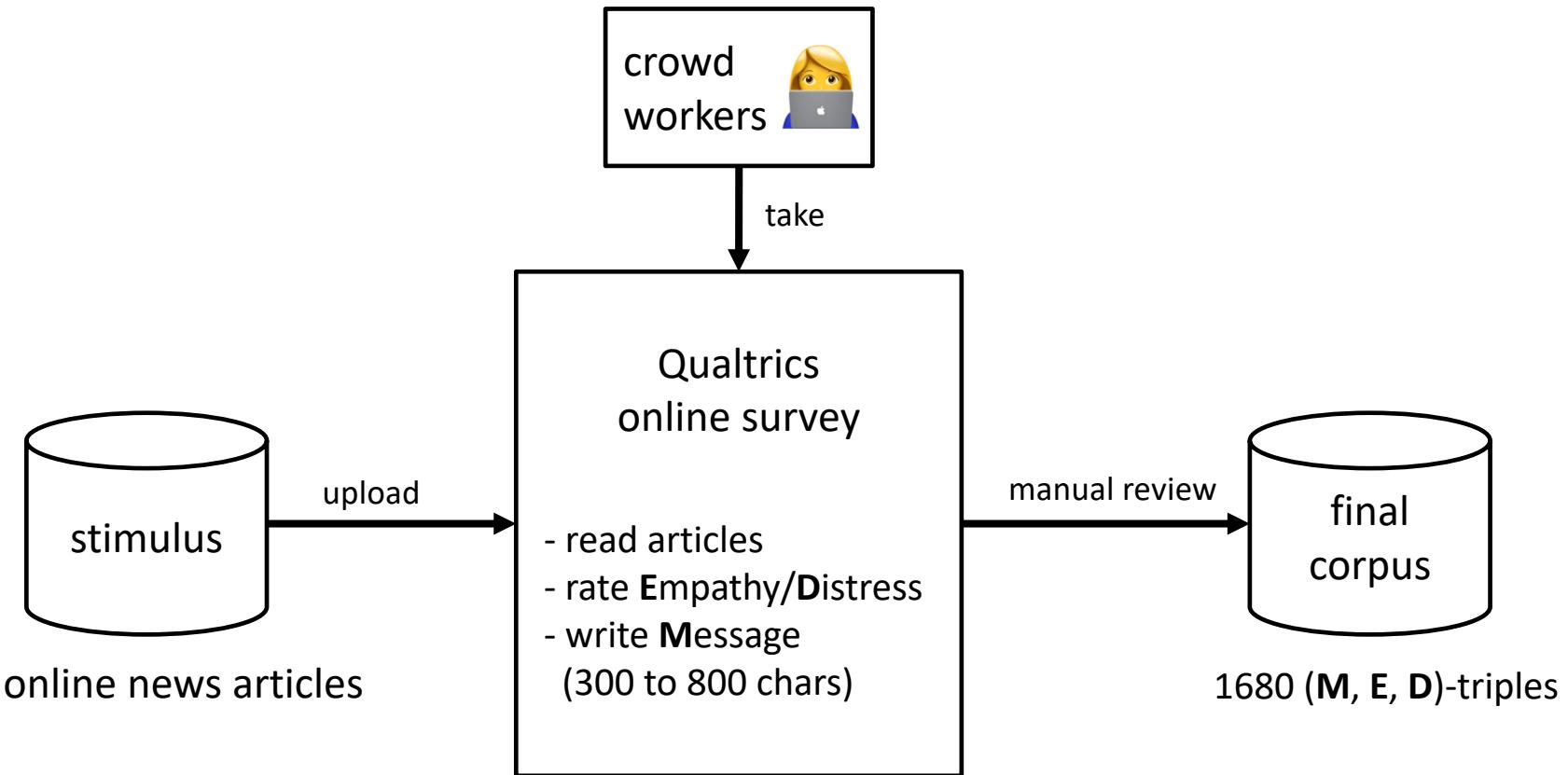
How strongly do you feel the following emotion? Using the 1-7 scale below, please indicate your agreement.



Corpus Creation Process



Corpus Creation Process



Split-half reliability around $r=.9$ for both empathy and distress 😊

Baseline Models for Empathy and Distress

- Models
 - Ridge Regression
 - Feed-Forward Net
 - Convolutional Neural Net (1 conv layer, filter sizes 1-3, 100 channels each)
 - RNN-type architectures did not work because of long sequences

Baseline Models for Empathy and Distress

- Models
 - Ridge Regression
 - Feed-Forward Net
 - Convolutional Neural Net (1 conv layer, filter sizes 1-3, 100 channels each)
 - RNN-type architectures did not work because of long sequences
- Features: FastText embeddings pre-trained on Common Crawl
- 10-fold cross-validation

Baseline Models for Empathy and Distress

- Models
 - Ridge Regression
 - Feed-Forward Net
 - Convolutional Neural Net (1 conv layer, filter sizes 1-3, 100 channels each)
 - RNN-type architectures did not work because of long sequences
- Features: FastText embeddings pre-trained on Common Crawl
- 10-fold cross-validation
- Correlation values around $r = .4$
- Ridge is viable but CNN significantly outperforms the others

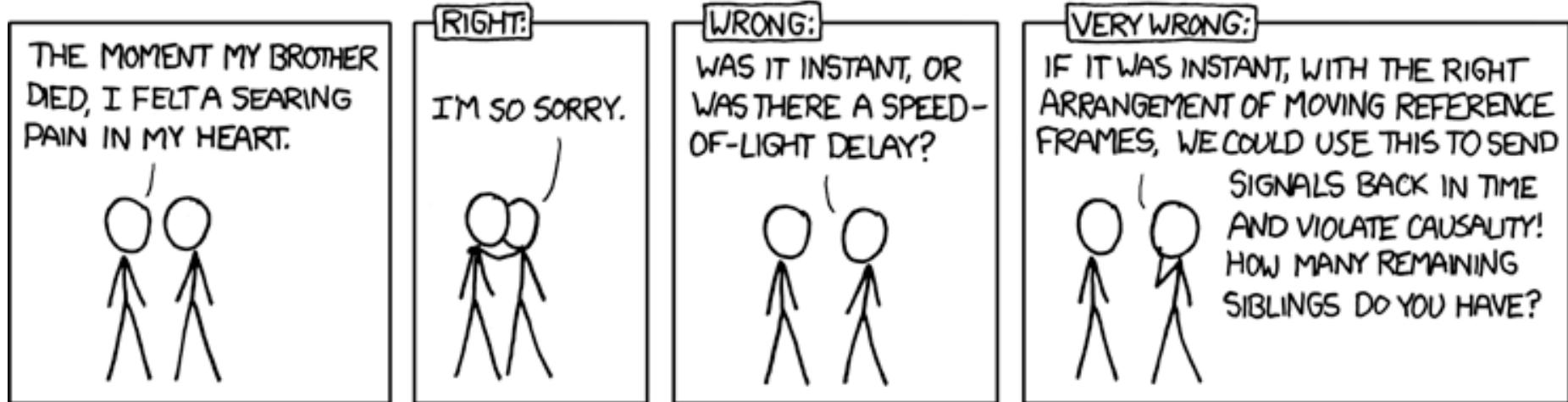
Conclusion

- Social media offers great opportunity to study empathy
- Modeling empathy received little attention for written language
- We presented the first publicly available gold standard
https://github.com/wwbp/empathic_reactions
- We distinguish Empathic Concern and Personal Distress
- New annotation methodology collects reliable ratings
from the experiencer using multi-item scales

Modeling Empathy and Distress in Reaction to News Stories

EMPATHY TIPS FOR ENGINEERS

~~SYMPATHY TIPS FOR PHYSICISTS~~



<https://xkcd.com/660/> CC-BY-NC 2.5

Sven Buechel *

Anneke Buffone *

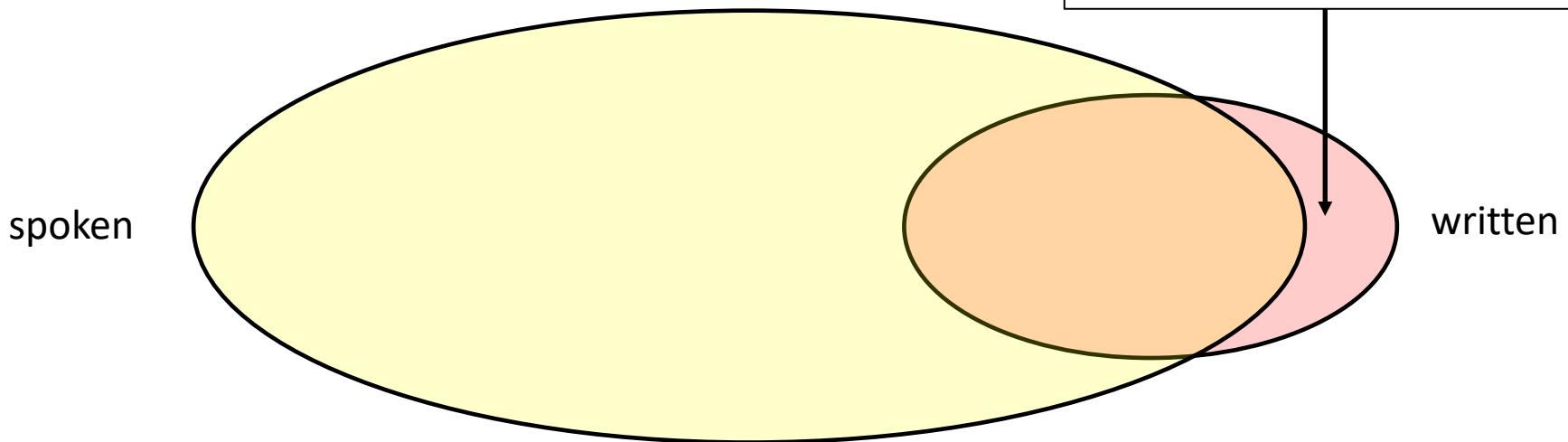
Barry Slaff

Lyle Ungar

João Sedoc

Backup Slides

Previous Work



Exemplary Entries

Empathy	Distress	Message
4.8	3.1	<p><i>I'm sorry to hear about Dakota's parents. No one wants that to happen and it's unfortunate that her parents couldn't work it out. I hope they are able to still remain civil around the kids and family. [...]</i></p>
4.0	5.5	<p><i>Here's an article about [a] crazed person who murdered two unfortunate women overseas. Life is crazy. [...] It feels like there's on place safe in this world to be a woman sometimes.</i></p>
1.0	1.3	<p><i>I just read an article about some chowder-head who used a hammer and a pick ax to destroy Donald Trump's star on the Hollywood walk of fame. [...] Lol, can you believe this garbage? Who has such a hollow and pathetic life that they don't have anything better to do with their time than commit petty vandalism because they dislike some politician? [...]</i></p>

Split-Half Reliability (SHR)

- Based on Pearson correlation r
- Flexible (works with crowdsourcing and best-worst scaling)
- Most popular in psychology
- Increasingly popular within CL (Mohammad et al.)

	i1	i2	i3	i4	i5	i6
d1						
d2						
d3						
d4						
d5						
d6						

Split-Half Reliability (SHR)

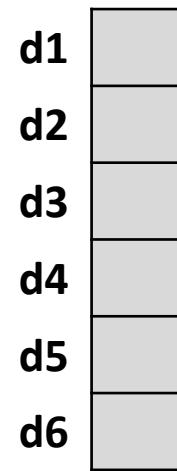
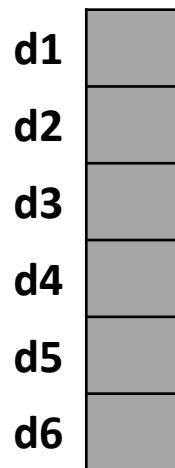
- Based on Pearson correlation r
- Flexible (works with crowdsourcing and best-worst scaling)
- Most popular in psychology
- Increasingly popular within CL (Mohammad et al.)

	i1	i4	i5
d1			
d2			
d3			
d4			
d5			
d6			

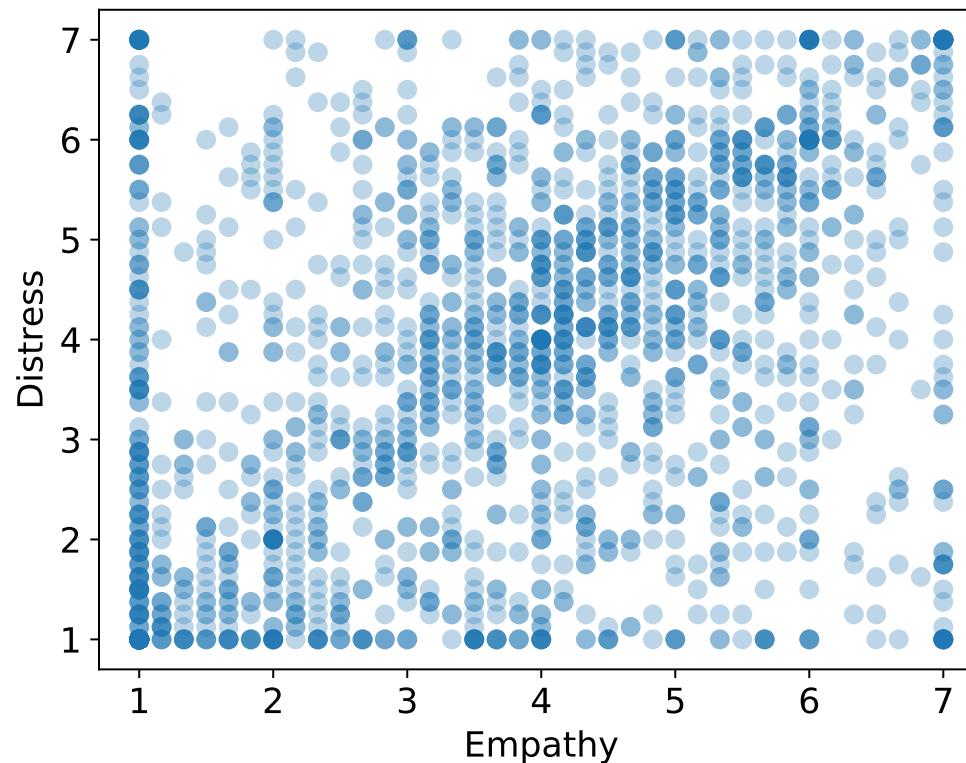
	i2	i3	i6
d1			
d2			
d3			
d4			
d5			
d6			

Split-Half Reliability (SHR)

- Based on Pearson correlation r
- Flexible (works with crowdsourcing and best-worst scaling)
- Most popular in psychology
- Increasingly popular within CL (Mohammad et al.)



Bivariate Rating Distribution



- Good coverage of the full range of rating scales
- Empathy and distress are distinct (moderate correlation of $r=.45$)

Models

- Ridge Regression
- Feed-Forward Net
 - two layers, 256 and 128 units
- Convolutional Neural Net
 - one conv layer
 - filter sizes 1, 2, 3
 - 100 output channels
 - average pooling
 - dense layer (128)

Experimental Setup

- Features: FastText embeddings pre-trained on Common Crawl
- Train distinct models for empathy and distress
- Exclude 10% of data for dev experiments
- 10-fold CV on remaining data
- Evaluate with Pearson correlation r

Results

	Empathy	Distress	Mean
Ridge	.385	.410	.398
FFN	.379	.401	.390
CNN	.404*	.444*	.424*

- Ridge regression is viable option, outperforms FFN
- CNN significantly outperforms Ridge and FFN
(* two-tailed paired t-test; $p < .05$)