Inducing Positive Perspectives with Text Reframing

Caleb Ziems,† Minzhi Li,† Anthony Zhang, Diyi Yang
Benefits of **Positivity**

- **Individuals**
  - Emotional well-being
  - Psychological well-being
Benefits of **Positivity**

- **Individuals**
  - Emotional well-being
  - Psychological well-being

- **Dyads**
  - Motivational Support
  - Negotiation
  - Persuasion
Benefits of Positivity

- **Individuals**
  - Emotional well-being
  - Psychological well-being

- **Dyads**
  - Motivational Support
  - Negotiation
  - Persuasion

- **Teams**
  - Reduce conflict
  - Increase collaboration
Benefits of **Positivity**

- **Individuals**
  - Emotional well-being
  - Psychological well-being

- **Dyads**
  - Motivational Support
  - Negotiation
  - Persuasion

- **Teams**
  - Reduce conflict
  - Increase collaboration

- **Corporations**
  - Branding
  - Customer relations
Benefits of **Positivity**

- **Individuals**
  - Emotional well-being
  - Psychological well-being

- **Dyads**
  - Motivational Support
  - Negotiation
  - Persuasion

- **Teams**
  - Reduce conflict
  - Increase collaboration

- **Corporations**
  - Branding
  - Customer relations

*[Caution: Avoid “fake” positivity]*
Inducing Positive Perspectives with Text Reframing

**Motivation:** machines can help people achieve these *positive benefits*
Inducing Positive Perspectives with Text Reframing

Motivation: machines can help people achieve these positive benefits

restructure thought
Inducing Positive Perspectives with Text Reframing

Motivation: machines can help people achieve these positive benefits

- **Restructure thought**

  (a) **Remove**: de-emphasize exaggerated forms of negativity
Inducing Positive Perspectives with Text Reframing

Motivation: machines can help people achieve these positive benefits

- **restructure thought**
  - (a) **remove**: de-emphasize exaggerated forms of negativity
  - (b) **reweight**: place more emphasis on positive elements with *new perspectives*
Inducing Positive Perspectives with Text Reframing

Motivation: machines can help people achieve these positive benefits

- **restructure thought**
  - (a) **remove**: de-emphasize exaggerated forms of negativity
  - (b) **reweight**: place more emphasis on positive elements with new perspectives
  - (c) **rephrase**: maintain fluency and genuineness without contradicting the original idea
Inducing Positive Perspectives with Text Reframing

**Motivation:** machines can help people achieve these positive benefits

- **restructure thought**
  - (a) **remove:** de-emphasize exaggerated forms of negativity
  - (b) **reweight:** place more emphasis on positive elements with new perspectives
  - (c) **rephrase:** maintain fluency and genuineness without contradicting the original idea

→ “positive reframing”
I absolutely hate making decisions.

I find myself having a lot of those to make.
I absolutely hate making decisions.
I find myself having a lot of those to make.

I absolutely love making decisions.
I find myself having a lot of those to make.

style transfer
I absolutely hate making decisions.
I find myself having a lot of those to make.

I absolutely love making decisions.
I find myself having a lot of those to make.
I absolutely hate making decisions. I find myself having a lot of those to make.

I absolutely love making decisions. I find myself having a lot of those to make.
I absolutely hate making decisions. I find myself having a lot of those to make.

Although I have a lot of decisions to make, it'll become easier once I start to get used to it.

Positive Reframe
I absolutely hate making decisions. I find myself having a lot of those to make.

Although I have a lot of decisions to make, it'll become easier once I start to get used to it.

I absolutely love making decisions. I find myself having a lot of those to make.
Ugh, it is SUCH a hot day 😢

Reframe:

This hot weather makes it the perfect day for ice cream 🍦
1. Building the Positive Psychology Corpus

1: Use distant supervision to identify negative posts
1. Building the Positive Psychology Corpus

1: Use *distant supervision* to identify negative posts

2: Apply Content Filters
1. Building the **Positive Psychology Corpus**

1: Use *distant supervision* to identify negative posts

2: Apply Content Filters

3: Data Annotation
1. **Building the Positive Psychology Corpus**

| Strategy        | ![Image](image) |
|-----------------|----------------|
| Growth Mindset  | ![Image](image) |
| Impermanence    | ![Image](image) |
| Neutralizing    | ![Image](image) |
| Optimism        | ![Image](image) |
| Self-Affirmation| ![Image](image) |
| Thankfulness    | ![Image](image) |
1. Building the **Positive Psychology Corpus**

| Strategy       | Distribution |
|----------------|--------------|
| Growth Mindset | 25.4%        |
| Impermanence   | 19.5%        |
| Neutralizing   | 36.1%        |
| Optimism       | 48.7%        |
| Self-Affirmation | 10.1%    |
| Thankfulness   | 13.0%        |
## 1. Building the Positive Psychology Corpus

| Genuineness | Strategy       | Distribution |
|-------------|----------------|--------------|
| 3.77 / 5.0  | Growth Mindset | 25.4%        |
| 4.03        | Impermanence   | 19.5%        |
| 3.53        | Neutralizing   | 36.1%        |
| 3.89        | Optimism       | 48.7%        |
| 3.75        | Self-Affirmation | 10.1%      |
| 3.95        | Thankfulness   | 13.0%        |
## 1. Building the Positive Psychology Corpus

| Agreement | Genuineness | Strategy       | Distribution |
|-----------|-------------|----------------|--------------|
| ICC = 0.59| 3.77 / 5.0  | Growth Mindset | 25.4%        |
| 0.60      | 4.03        | Impermanence   | 19.5%        |
| 0.32      | 3.53        | Neutralizing   | 36.1%        |
| 0.44      | 3.89        | Optimism       | 48.7%        |
| 0.42      | 3.75        | Self-Affirmation| 10.1%       |
| 0.42      | 3.95        | Thankfulness   | 13.0%        |
2. Automatic Positive Reframing
2. Automatic Positive Reframing

Ugh, it is SUCH a hot day 😫

Reframe:

This hot weather makes it the perfect day for ice cream 🍦
2. Automatic Positive Reframing

Objective: \[ \frac{1}{N} \sum_{i=0}^{N} \log p(g_i | g_{0:i-1}) \]

\[ g = \{ s, \psi_t, t \} \]

Sometimes I get these impulses to just throw a tantrum. Like throw/break things, cry and scream.

Source
2. Automatic Positive Reframing

**Objective:**
\[
\frac{1}{N} \sum_{i=0}^{N} \log p(g_i | g_{0:i-1})
\]

\[g = \{ s, \psi_t, t \} \]

Sometimes I get these impulses to just throw a tantrum. Like throw/break things, cry and scream.

**Source**

**self-affirmation**

**Reframe Strategies**
2. **Automatic Positive Reframing**

**Objective:**
\[
\frac{1}{N} \sum_{i=0}^{N} \log p(g_i|g_{0:i-1})
\]

\[g = \{s, \psi_t, t\}\]

\[= g_1, g_2, \ldots, g_K, g_{K+1}, g_{K+2}, \ldots, g_{K+L+1}, g_{K+L+2}, \ldots, g_{K+L+M+2}, g_N\]

\[s_1, s_2, \ldots, s_K, [\text{strg}], \psi_1, \ldots, \psi_L, [\text{refr}], t_1, \ldots, t_M, [\text{EOS}]\]

- **Source:** Sometimes I get these impulses to just throw a tantrum. Like throw/break things, cry and scream.
- **self-affirmation:** It's normal for to feel overwhelmed sometimes but I know I am strong to handle and go through it.
- **Reframe Strategies**
- **Target**
2. Positive Reframing

\[ p(t|s) \textbf{Unconstrained:} \quad \text{Human evaluation} \]

- Meaning
2. Positive Reframing

\[ p(t|s) \text{ **Unconstrained:**} \quad \text{Human evaluation} \]

- Meaning
- Positivity
2. Positive Reframing

\[ p(t|s) \textbf{Unconstrained:} \quad \text{Human evaluation} \]

- Meaning
- Positivity
- Fluency
2. Positive Reframing

\( p(t | s) \) **Unconstrained:** Human evaluation

- **Meaning**
- **Positivity**
- **Fluency**

BART
2. Positive Reframing

\[ p(t|s) \text{ Unconstrained:} \quad \text{Human evaluation} \]

- **Meaning**
- **Positivity**
- **Fluency**

![Graph showing comparison of T5 and BART models on Meaning, Positivity, and Fluency dimensions with human evaluation.](image)
2. Positive Reframing

$p(t \mid s)$ Unconstrained: Human evaluation

- **Meaning**
- **Positivity**
- **Fluency**

| Model   | Meaning | Positivity | Fluency |
|---------|---------|------------|---------|
| seq2seq LSTM | ![seq2seq LSTM Meaning](image) | ![seq2seq LSTM Positivity](image) | ![seq2seq LSTM Fluency](image) |
| T5      | ![T5 Meaning](image) | ![T5 Positivity](image) | ![T5 Fluency](image) |
| BART    | ![BART Meaning](image) | ![BART Positivity](image) | ![BART Fluency](image) |
2. Positive Reframing

\[ p(t|s) \textbf{Unconstrained:} \] Human evaluation

- **Meaning**
- **Positivity**
- **Fluency**

![Bar chart comparing GPT-2, seq2seq LSTM, T5, and BART on Meaning, Positivity, and Fluency](chart.png)
2. Positive Reframing

\[ p(t|s) \textbf{Unconstrained:} \]

Meaning

|Model| 1 | 2 | 3 | 4 |
|-----|---|---|---|---|
|CopyNMT|   |   |   |   |
|GPT-2|   |   |   |   |
|seq2seq LSTM|   |   |   |   |
|T5|   |   |   |   |
|BART|   |   |   |   |

Positivity

Fluency

Human evaluation
2. Positive Reframing

\( p(t|s) \) **Unconstrained:** Human evaluation

| Model       | Meaning | Positivity | Fluency |
|-------------|---------|------------|---------|
| BART (Unconstrained) | 4.13    | 3.81       | 4.15    |
2. Positive Reframing

\[ p(t|s, \psi_t) \] **Controlled:**

**Performance boost**
from targeting gold strategies

| Model            | Meaning | Positivity | Fluency |
|------------------|---------|------------|---------|
| BART (Unconstrained) | 4.13    | 3.81       | 4.15    |
| BART (Controlled)  | 4.23    | 4.07       | 4.27    |
2. Positive Reframing

**Few Shot:**

Loss in *meaning* preservation, despite increase in *positivity*.

| Model                  | Meaning | Positivity | Fluency |
|------------------------|---------|------------|---------|
| BART (Unconstrained)   | 4.13    | 3.81       | 4.15    |
| BART (Controlled)      | 4.23    | 4.07       | 4.27    |
| GPT-3 (5-shot)         | 3.73    | 4.17       | 4.27    |
I just went back to school today and I'm already stressed because we have midterms next week and this weekend – JAM PACKED :-(

Well, just think by the end of next week we'll be all done with the midterms!

I just went back to school today and I'm already stressed because we have midterms next week and this weekend – JAM PACKED
2. Positive Reframing: Example Reframes

Could someone just give me like $1000? It would change my life, stressing about rent, bills and food money is just the worst.

Time to focus on making a budget, so I don’t have to stress about rent, bills, and food money.

I need to learn how to manage my money better so that I don’t have to stress about rent, bills and food money.
So glad that tomorrow is Friday. This has seriously been the longest week of my life! 

tired

I'm glad the weekend is coming up, so I can rest.

This week has been long, but I'm looking forward to the weekend.

2. Positive Reframing: Example Reframes
Sometimes I get these impulses to just throw a tantrum. Like throw/break things, cry and scream.

It's normal for to feel overwhelmed sometimes but I know I am strong to handle and go through it.

Sometimes I get these impulses to just throw a tantrum. Like throw/break things, cry and scream. But I'm strong, and I know I can handle it.
Conclusion

1. *Introduced:* New NLG Task
Conclusion

1. *Introduced:* **New NLG Task**

2. *Built:* **Positive Psychology Frames**
Conclusion

1. Introduced: New NLG Task

2. Built: **Positive Psychology Frames**

3. Tested: Baseline Style Transfer Models for **Automatic Positive Reframing**
Conclusion

1. **Introduced:** New NLG Task (Far from Solved!)

2. **Built:** POSITIVE PSYCHOLOGY FRAMES

3. **Tested:** Baseline Style Transfer Models for **Automatic Positive Reframing**

   - **Future Work:** enforce semantic equivalence + allow new perspectives
     - Semantic Attention-Guided Decoding (Krause et al. 2020)
     - Pragmatic Reconstruction (Nie et al. 2020)
     - Persona Consistency (Kim et al. 2020)
Inducing Positive Perspectives with Text Reframing

Caleb Ziems, * Minzhi Li, * Anthony Zhang, Diyi Yang

GT-SALT/positive-frames
talacre beach really need to sort their checking in routine. Caravans are never ready 30 minutes later and still sat in the car waiting..

Talacre beach's check-in process could do with some improvement to reduce the wait time for customers.

The caravans at Talacre beach need to sort out their check-in process, so the kids don't get stuck in the car.
The fact that my mom's been calling me every night to calm me down from school says so much. I've needed it these past few days.

My mom has been calling me every night to calm me down from school. I've needed it these past few days. I'm thankful for her.

I'm really lucky to have such a caring mum who is willing to call me every night when I'm having a hard time.
2. Automatic Positive Reframing

1. **NEGATIVE:** "I am the only person I know who writes a healthy grocery list and plans meals when I am stressed: (Can'tSleep)"
   **POSITIVE:** "I'm so thankful that I am still able to eat healthy even when I'm stressed."

2. **NEGATIVE:** "I am SO stressed with all my exams and my lit review hanging over my head this week."
   **POSITIVE:** "Only one more week until my exams and lit review are all done!"

3. **NEGATIVE:** "Ugh my mac is starting to slow up and I need to figure out how to defragment the hard drive..."
   **POSITIVE:** "I need to defragment the hard drive to speed up my mac. Good thing I’m smart, and I know I can do this."

4. **NEGATIVE:** "This has been like the worst week ever im so done with everything. sick tired"
   **POSITIVE:** "I made it to the end of the most challenging week ever!"

5. **NEGATIVE:** "I have a huge project due tomorrow morning. But where do I have to be, a stupid basketball game dumb"
   **POSITIVE:** "I should plan ahead next time so that my basketball game does not conflict too closely with my projects."
2. Positive Reframing: Error Analysis

1. insubstantial changes (26%)
   (especially for neutralizing category)

2. contradictions to the premise (9%)

3. self-contradictions (6%)

4. hallucinations (2%)
2. Positive Reframing

**ROUGE:**  **R-1**  **R-2**  **R-L**  
(meaning)

**BERT Score:**  **BScore**  
(meaning)

**Positivity Delta:**  **Δ TB**  
(sentiment)

Meaning  Positivity  Fluency
(on 5-point Likert scale)

|                      | Model                  | Human Evaluation |
|----------------------|------------------------|------------------|
|                      |                        | Meaning | Positivity | Fluency |
| Few-shot             | GPT-3                  |         |            |
|                      | GPT-Neo                |         |            |
|                      | GPT                    |         |            |
|                      | GPT-2 No-pretrain      |         |            |
|                      | GPT-2                  |         |            |
|                      | Seq2Seq-LSTM           |         |            |
|                      | CopyNMT                |         |            |
|                      | T5                     |         |            |
|                      | BART                   |         |            |
|                      | **T5**                 | 3.80    | 3.82       | 4.18    |
|                      | **BART**               |         |            |
2. Positive Reframing

**Few-Shot:**

- *High* Positivity and Fluency
- *Less* Meaning preservation

| Model       | Meaning | Positivity | Fluency |
|-------------|---------|------------|---------|
| Few-shot GPT-3 | 3.73    | 4.17       | 4.27    |
| GPT-Neo      | 3.69    | 4.16       | 4.21    |
| GPT          |         |            |         |
| GPT-2 No-pretrain |       |            |         |
| GPT-2        |         |            |         |
| Seq2Seq-LSTM |         |            |         |
| CopyNMT      |         |            |         |
| T5           |         |            |         |
| BART         |         |            |         |
| T5           |         |            |         |
| BART         |         |            |         |
| Human        | 3.80    | 3.82       | 4.18    |
2. Positive Reframing

\[ p(t \mid s) \text{ Unconstrained:} \]

Reasonable **Positivity** and
Fluency **while better preserving**
Meaning

| Model | Human Evaluation |
|-------|------------------|
|       | Meaning | Positivity | Fluency |
| Few-shot |        |            |         |
| GPT-3   | 3.73    | 4.17       | 4.27    |
| GPT-Neo | 3.69    | 4.16       | 4.21    |
| GPT     | 3.55    | 3.91       | 4.08    |
| GPT-2 No-pretrain | 3.11 | 3.66       | 3.96    |
| GPT-2   | 3.58    | 4.01       | **4.18**|
| Seq2Seq-LSTM | 3.33 | 4.15       | 4.10    |
| CopyNMT | 3.57    | 3.69       | 3.91    |
| T5      | 4.09    | 3.79       | 4.06    |
| BART    | **4.13**| **3.81**   | 4.15    |
| T5      |         |            |         |
| BART    |         |            |         |

| Human | 3.80 | 3.82 | **4.18** |
2. **Positive Reframing**

\[ p(t, \psi_t|s) \] **Prediction:**

Performance remains stable

| Model                  | Human Evaluation |
|------------------------|------------------|
|                        | Meaning | Positivity | Fluency |
| Few-shot               |         |            |         |
| GPT-3                  | 3.73    | 4.17       | 4.27    |
| GPT-Neo                | 3.69    | 4.16       | 4.21    |
| GPT                    | 3.55    | 3.91       | 4.08    |
| GPT-2 No-pretrain      | 3.11    | 3.66       | 3.96    |
| GPT-2                  | 3.58    | 4.01       | 4.18    |
| Seq2Seq-LSTM           | 3.33    | 4.15       | 4.10    |
| CopyNMT                | 3.57    | 3.69       | 3.91    |
| T5                     | 4.09    | 3.79       | 4.06    |
| BART                   | 4.13    | 3.81       | 4.15    |
| T5                     | 4.10    | 3.64       | 4.11    |
| BART                   | 4.09    | 3.95       | 4.11    |
| Human                  | 3.80    | 3.82       | 4.18    |
2. Positive Reframing

\[ p(t | s, \psi_t) \] Controlled:

Performance boost from targeting gold strategies
2. **Positive Reframing**

\[ p(t|s, \psi_t) \text{ Controlled: } \]

*Performance boost from targeting gold strategies*