

## **Should climate scientists fly? A case study of arguments at the system level**

Forthcoming, *Informal Logic*

Jean Goodwin

### **Appendix 1: Data and Methods**

#### **Corpus 1: Tweets**

Corpus 1 consists primarily of Twitter original posts from January 2010-April 2020 containing the keywords “climate” and “fly”, plus retweets and replies. It was collected through Crimson Hexagon, a social media analytics platform giving access to the historical Twitter firehose. The full corpus included approximately 341K entries. The following subcorpora were created:

- SCIENTISTS: original posts and replies from Corpus 1 including keyword “scient\*”; 4.5K tweets, 654 of which were relevant to the project. These were coded for standpoint: HYPOCRISY claim, RESPONSE to hypocrisy claim, discussion of FLYLESS. Tweets expressing the HYPOCRISY theme were checked against their original conversations to determine whether they were from climate skeptics, climate believers, or undetermined, and whether they were by scientists, nonscientists, or undetermined.
- HYPOCRISY: original posts and replies from Corpus 1 including keyword “hypocr\*”; 4.6K tweets. These were scanned for relevance to climate scientists flying.
- METADISCOURSE 1: original posts and replies from Corpus 1 including keywords related to argumentation, including terms that were found in Corpora 2 and 3: “argu\*, fallac\*, logic\*, illogic\*, ad hominem, ad hom, [no true] Scotsman, [non] sequiter, poisoning [the well], red herring, special pleading, straw man, [tu] quoque, whatabout\*”; 544 tweets..

After exploration, Corpus 1 was supplemented with an additional search:

- METADISCOURSE 2: Original posts and replies from January 2010-April 2020 including the keyword climate and hypocrisy terminology (“hypocrisy, hypocrite, hypocritical”) within 5 words of argumentation terminology (“argument, line, point, narrative, fallacy, idea”); 108 tweets. These were scanned for relevance.

#### **Corpus 2: Longform**

Corpus 2 consists of journalism, blog posts and comments, podcasts, and any other material available on the internet. This corpus was collected through:

- (a) The top 100 URLs referenced in each year in Corpus 1 from 2017 (when the data became available) to 2020, plus the top 100 URLs reference in the SCIENTISTS subcorpus. These were scanned for relevance.
- (b) Climate skeptics were underrepresented in (a), so Google search for scientists hypocrisy or flying was performed on skeptical blogs Climate Audit, Watts Up With That, and Climate Depot.
- (c) Convenience sample: material I had been collecting over several years of interest in this topic.
- (d) Snowball sample: Other longform mentioned in material in Corpora 1, 2 or 3.

There are 112 documents in this corpus. The documents were imported into Atlas.ti for analysis. The corpus was first holistically coded (Saldaña 2016) for standpoint in debate. Provisional coding was then applied to the non-skeptics' discourse identify main lines of argument I expected, based on several years acquaintance with the controversy over climate scientists' flying and the existing literature. The code set was constantly updated in the process to add new arguments, to split codes applied to heterogeneous discourse, and to ensure symmetry between the standpoints (i.e, to actively look for responses to each side's major arguments). Repeated passes through the data were made as the code set evolved. In vivo coding was also used to capture argument metadiscourse. Direct interactions between individuals with different standpoints were flagged, along with interesting material that could not be captured within the coding scheme. A second cycle of coding adopted a more open coding approach to differentiate themes within the existing codes.

### **Corpus 3: Twitter Conversations**

Corpus 3 consists of sets of tweets replying to an original post, ranging from 3 to hundreds. This corpus was collected through:

- (a) Locating and checking for relevance the conversations in which tweets in Corpus 1 had originally appeared.
- (b) Snowball sample: Further Twitter conversations mentioned (e.g., quote-retweeted) in (a).
- (c) Convenience sample: Twitter conversations I had been collecting over several years of interest in this topic.

There are 85 conversations in this corpus. Conversation texts were reconstructed using Treeverse and imported into Atlas.ti for analysis, with the exception of several conversations which were too large for capture. They was coded through the provisional and in vivo approaches described above. Finally, the analysis of the FlyLess arguments in all three corpora were checked by posting a summary on Twitter in conversation with several scientists whose discourse is in the corpus.