# Writing Workshop Day 3

**2. Identify thematic strings in a paragraph from an example paper**

"Even if you use Twitter only for professional purposes, consider opening up a little bit to show your followers you are a real person. People outside your field are not likely to follow you if your tweets are only about sharing events, articles, and positions in your own field. You need to add an extra ingredient—your opinions, or something personal—to what you share. One way to do this is through sharing failures: a rejected paper or job application, or even a spilt coffee. This is a great way to give and receive moral support from other academics."

Make two lists. In one, list the characters you remember. In the other, list just two or three words that would capture the central concepts that the writer weaves around those characters, words that constitute the conceptual center of that paragraph.

1. Characters:
    - the twitter user (the reader), their followers, Twitter
2. Concepts:
    - humanity, relatable, friendly

Go back and look at the paragraph. Do the thematic strings in the paragraph agree with the characters and concepts that you remembered and listed?

I found that this paragraph was easy to understand when I first read it. I remembered characters and concepts from the paragraph that, after reading again, agree with its thematic strings. To me, the thematic strings revovle around being personable on your Twitter and not only posting around detailed scientific information.

**3. Identify thematic strings in a paragraph from your manuscript.**

Pick a paragraph from your manuscript that you think "hangs together". Do the same exercise for this paragraph and identify the thematic strings.

Next pick a paragraph from your manuscript that feels vague when you re-read it. Do the same exercise, and see if you can revise the paragraph so that its thematic strings are consistent and clear.

Hangs together:

Another area of concern in the NOAA storm events database arises with geographic bias. This bias explains the idea that there may be inconsistencies in the recording of events due to geographic location. First, there may be differences in the supply of information from different regions based on population differences or locations of weather stations. Secondly, structural changes in the database create differences in how location data is obtained and recorded.

**Refocus subjects, add in topic strings, reorder sentences**

1. Characters
    - database, recorded events
2. Concepts
    - geography, location, bias/differences, data entry

Vague:

Some temporal changes are structural, in terms of the database changing what types of events are recorded.

In the NOAA database, from 1950 to 1954, only tornado events were recorded [@stormeventsdetails]. Then, from 1955 to 1996, only tornado, thunderstorm wind and hail events were recorded as digital data [@stormeventsdetails]. Starting in 1996, the database began including all 48 event types that are recorded today [@stormeventsdetails]. Additionally, the NWS reports that there was reduced funding from 1976 to 1979 that led to a decrease in flood damage data collection [@downton2005reanalysis]. Then, after 1983, Congress began to require The US Army Corps of Engineers to provide yearly flood damage reports in contract with the NWS [@downton2005reanalysis]. This provides a more consistent report of flood damage that may not have otherwise been reported. These changes pose inconsistensies in the number of certain storm events over large periods of time.

1. Characters
    - the database, temporal changes, NWS, Congress
2. Concepts
    - changes, time, data entry
    - seems like there is a lot going on in this paragrpah even though I can tell its about changes over time

Revised:

Some temporal changes are structural, in terms of the database changing what types of events are recorded. Since 1950, the database has undergone two major changes in event recording. The first change occured in 1954 when the NOAA database starting recording thunderstorm wind and hail in addition to the tornado events it had already been recording [@stormeventsdetails].Fourty-two years later, the database started recording all 48 event types that it continues to record today [@stormeventsdetails].

These major changes were also influenced by smaller changes in funding and policy. One of these changes occured between 1976 and 1979 when the NWS received reduced funding from the US government that decreased flood damage data collection [@downton2005reanalysis]. This decreased flood data collection may have produced an artificially low number of recorded floods during that time. To counteract this, in 1983, Congress began to require The US Army Corps of Engineers to provide yearly flood damage reports in contract with the NWS [@downton2005reanalysis]. This could have the opposite effect in that the data may provide a more consistent report of flood damage that would not have otherwise been reported. These changes pose inconsistensies in the number of certain storm events over large periods of time.

Maybe I could make a table with the dates instead of trying to list them in paragraph form??

**4. Identify the issue in paragraphs from example papers.**

"Even if you use Twitter only for professional purposes, consider opening up a little bit to show your followers you are a real person."

What theme is introduced at the end of this issue?

- opening up as a real person on social media

Does the rest of the paragraph follow through in talking about this theme?

- Yes because the rest of the paragraph gives examples of how you can do this

From the current version of the issue sentence, is it easy to pick out the main concept that will be developed in the rest of the paragraph (the discussion)?

- I think that the rest of the paragraph does follow through with the issue and it is easy to tell what this issue is from the first sentence.

Next, re-read the following paragraph from the Gall et al. paper:

"**A more subtle form of introducing hazard bias arises from issues of the definition of the hazard and assigning loss estimates (by the original data source) to predefined hazard categories within a database.** This is most apparent in the management of complex events involving multiple hazards versus

a singular hazard event. A tornado spawned by a hurricane is counted as a unique tornado event, but it could also be lumped together within the entire hurricane event, or both. Each loss database classifies events differently, especially when they involve multiple hazard types (Guha-Sapir and Below 2002). Inconsistent naming conventions and classification methodologies aggravate this problem and can result in different (and/ or artificial) hazard categories for similar, if not identical events. For example, Downton et al. (2005) reveal a $520 million"flood" loss in FEMA's database that was not in the NWS data. The discrepancy is a result of differences in how each agency defines what constitutes a flood event. In this case, the event (storm surge) was outside NWS's definition of a flood."

Highlight the text that gives the issue in this paragraph. What theme is identified at the end of the issue sentence?

- assigning loss estimates to categories

Does the rest of the paragraph follow through on this theme?

- Yes mostly. It gives an example of how a hurricane event and its effects would be categorized. However, then it starts also talking about different hazard categories.

In the following version of the paragraph, I have rewritten the first sentence.

"More subtly, hazard bias arises because databases differ in how they define disasters and sort hazards into predefined categories. This is most apparent in the management of complex events involving multiple hazards versus a singular hazard event. A tornado spawned by a hurricane is counted as a unique tornado event, but it could also be lumped together within the entire hurricane event, or both. Each loss database classifies events differently, especially when they involve multiple hazard types (Guha-Sapir and Below 2002). Inconsistent naming conventions and classification methodologies aggravate this problem and can result in different (and/ or artificial) hazard categories for similar, if not identical events. For example, Downton et al. (2005) reveal a $520 million"flood" loss in FEMA's database that was not in the NWS data. The discrepancy is a result of differences in how each agency defines what constitutes a flood event. In this case, the event (storm surge) was outside NWS's definition of a flood."

Does this version of the paragraph seem more or less coherent that the original version? Identify the differences in the structure of the two versions of the sentence that contribute to coherence of the paragraph.

- This version seems more coherent than the original. This sentence is more clear because it covers both issues of defining event types and separating larger hazard categories.

**5. Identify the issue in paragraphs from your manuscript.**

Pick a series of four or five paragraphs in your manuscript. For each paragraph, highlight the sentence that states the issue. What concept is presented at the end of that sentence? Highlight places in the rest of that paragraph that follow up on that theme. Are there any thematic strings in the rest of the paragraph that weren't "announced" in the issue sentence?

Re-write the issue sentence of each paragraph. As you do, keep in mind the advice from the "Sentences" section on characters and actions. Also, make sure that you place the theme you want to develop at the end of the issue sentence.

## Geographic Bias

### Introduction

Another area of concern in the NOAA storm events database arises with geographic bias. **This bias explains the idea that there may be inconsistencies in the recording of events due to geographic location**. First, there may be differences in the supply of information from *different regions* based on population

differences or locations of weather stations. Secondly, structural changes in the database create differences in how *location data* is obtained and recorded.

Revised:
Another area of concern in the NOAA storm events database arises with geographic bias which explains the idea that there may be **inconsistencies in the recording of events due to geographic location**.

### Structural

**There can also be changes at country or state level over time that lead to excluding or double counting events or loss data** [@gall2009losses]. The *NWS changed its reporting strategy* from loss estimates by climate region to loss estimates in specific counties where event occurred around 1995 [@gall2009losses]. The NOAA website provides a disclaimer that "the source data ingested into the database are widely varied and leads to many *questions about the precision and accuracy of the location data*" [@stormdatafaq].

Revised:

Geographic bias also results from structural changes at the county or state level. In 1995, the NWS changed its reporting strategy from loss estimates by climate region to loss estimates in specific counties where the event occurred [@gall2009losses]. This type of change could lead to excluding loss data or double counting events [@gall2009losses].

### Aggregation of reporting

**Currently, the smallest unit of aggregation to use all parts of the database are Weather Forecasting Offices**. There are currently about 122 nationwide [@konisky2016extreme]. There may be inconsistencies between where county regions are defined and where *weather forecasting offices* are located. NOAA Storm Events states that "a hydrometeorological event will be referenced, minimally, to the nearest hundredth of a mile, to the geographical center (not from the village/city boundaries or limits) of a particular village/city, airport, inland lake, or location providing that the reference point is documented in the Storm Data software location database" [@nwsinstruction]. *Problems may arise here if the exact location of an event is not already documented* in the Storm Data software location database.

Revised:

When a location is entered for an event in the NOAA Storm Events database, inconsistencies may arise between where the event actually occured and the location that is recorded in the database. Currently, the smallest unit of aggregation to use all parts of the database are called Weather Forecasting Offices. Each office covers a geographic area of the US and there are currently about 122 nationwide [@konisky2016extreme]. These geographic areas overlap with county outlines but do not always directly coincide. It is unclear exactly how NOAA decides when to assign an event to a county or a Weather Forecasting Office. In regards to this decision, the NWS documentation states that "a hydrometeorological event will be referenced, minimally, to the nearest hundredth of a mile, to the geographical center (not from the village/city boundaries or limits) of a particular village/city, airport, inland lake, or location providing that the reference point is documented in the Storm Data software location database" [@nwsinstruction].

Some events are reported by forecast zone (`CZTYPE` of "Z") rather than county (`CZTYPE` of "C"). **Specific types of events are typically either always reported for a county or always reported for a forecast zone (see table below)**. Events typically reported by **county** include floods, tornado-like events, and a few other events often related to thunderstorms. Events typically reported by **forecast zone** include severe winter weather, extreme heat, events related to the water or coast, and a few others ("High Wind", "Dense Fog", "Strong Wind", "Wildfire", "Dust Storm", "Dense Smoke").

Revised: The NOAA Storm Events database also uses forecast zones as another unit of aggregation. Forecast zones are created as subsets of counties for more specific location data. Specific types of events are typically either always reported for a county (`CZTYPE` of "C") or always reported for a forecast zone (`CZTYPE` of "Z")

(see table below)**. Events typically reported by county include floods, tornado-like events, and a few other events often related to thunderstorms. Events typically reported by forecast zone include severe winter weather, extreme heat, events related to the water or coast, and a few others ("High Wind", "Dense Fog", "Strong Wind", "Wildfire", "Dust Storm", "Dense Smoke").

**This demonstrates the difficulty of narrowing down where exactly certain events occured.** It is important to keep in mind that some events might be very localized (for example, a flash flood, or even storm surge that's only near the coast of the county), but the county-level reporting of the event means that the whole county gets classified as experiencing or not experiencing the event.

### 6. Identify the issue in every paragraph of your manuscript.

Now go through your entire manuscript and highlight the issue sentence in each paragraph.

This should help highlight the structure of your paper and the major ideas you are trying to share. Do your issues agree with what your goals for this paper? Do they align with the "Challenge" at the end of your Introduction?

Make two lists: one with any issues that are in your paper but do not align with your Introduction's Challenge and one with any points that you were hoping to make in your manuscript but that aren't yet in the issue of a paragraph. Use these two lists to mark paragraphs that you might consider cutting (or revising) and to add "placeholders" for spots where you want to add paragraphs to address issues you've missed.

Issues that do not align:

- systemic bias section
- first few paragraphs of describing what the database is are choppy
- contradictory information about rip currents
- logarithmic scale for monetary losses

Points hoping to make:

- expand on inaccuracies of recording in terms of monetary losses for accounting bias
- defintiely a lot of room to expand in all of the bias sections
- population growth and changes
- changes in value of US dollar
- how to differentiate between Weather Forecasting Offices and forecast zones