TAGS: Tweet Dataset for Unitizing and Coding Decisions

(This activity is Adapted from Merrigan, Gerianne, and Carole Logan Huston. 2016. *Instructor's Manual and Test Bank to accompany Communication Research Methods*. 3rd Edition. New York: Oxford University Press.*)*

First, assign the students to code 20 tweets from the included TAGS data sheet without talking to anyone else. They should read through all of the tweets in the dataset and then determine 3 key terms or themes that they can use code the tweets. Alternatively, faculty member could provide 3-5 terms. Students should record a hash mark (if you print the sheet or copy the tweet into the column) for each unit of text that fits into each of the categories of the columns.

Next, pair students who coded their exact same set of tweets. Ask students to first, calculate unitizing reliability with the following formula:

Unitizing Reliability = (# of units you identify – # of units your partner identifies) ----------
---

(# of units you identify + # of units your partner identifies)

Second, each pair of students will need to calculate their inter-coder reliability using the following formula: Inter-Coder Reliability = # of times you and your partner AGREED on a category

Finally, have the student pairs record their percentages on the board (or use a table like this one):

	# unitizing decisions	# category decisions agreed exactly	Percentage of agreement
Unitizing decisions			
Categorizing tweets			
Categorizing media			
Categorizing thematic focus			

If the two coders cannot agree upon 70% or more of the decisions, they need to clarify operational definitions, or re-train the coders in order to achieve acceptable measurement reliability. Notice how many of the students achieved "acceptable" levels of agreements. If your students are doing their own projects, work on creating codebooks that they can use to code messages reliably for their studies. (Bonus Activity: Have the students make bar charts, pie charts or bar graphs that show the distribution of message categories as agreed on by the coders).