Opinion & Advocacy Standards for Social Media Measurement Validity Test Results & Recommendations

Meeting of the Conclave on Social Media Measurement Standards
Authors:

Julie O'Neil, Ph.D., Texas Christian University, Member IPR Measurement Commission
Michelle Hinson, CyberAlert; Chair of the IPR Measurement Commission
Katie Delahaye Paine, Paine Publishing, LLC, Member IPR Measurement Commission

Durham, NH, USA
October 14, 2016

Methodology

- Three organizations:
  - 1. Non-Profit – Goodwill Industries
  - 2. Consumer – Southwest Airlines
  - 3. Government – US Fish & Wildlife
- 8,510 clips collected
  - 204 clips
  - 201 posts analyzed k
- Sampling: Southwest and Goodwill had a random samples. USFWS was the actual number of qualified posts for the time period.
- All counts were as follows:
  - USFWS   Tweets 65   FB 74
  - Goodwill   Tweets 174   FB 312
  - Southwest   Tweets 178   FB 7,707
- 3-4 Coders on each round with no discussion between coders after coding begins
- Three test rounds

Standard Coding Instructions – Qualification of Items

To qualify for the study, each item had to meet the following criteria:

- Mention one of the three organizations

- Appear in an item dated between May 1, 2016 – July 31, 2016

- Must be "earned" i.e. neither paid media nor content created by or owned by the brand being studied.

- Straight retweets of organization-created tweets did not qualify for the study

- Modified retweets in which the author comments or expresses an opinion were coded.

- If clarification was needed coders followed link provided in tweet but only coded for the content in the 140 characters

- Pictures in posts were analyzed in the post to determine sentiment and messaging

STANDARDS TEST: Sentiment/Tone

► Sentiment should be coded based on the perspective of the average reader.

- Positive - An item leaves the reader more likely to support, recommend, and/or work or do business with the brand
- Negative - An item leaves the reader less likely to support, and/work or do business with the brand
- Neutral - An item contains no sentiment at all, just reports the facts.
- Balanced - contains equal amounts of positive and negative content"

► Coding social media posts, also involves looking at pictures in the post to determine sentiment and messaging.

STANDARDS TEST: Quality Measures

► Visual

  ▷ Inclusion of any chart, graph, logo, photo or other image that conveys a company 'presence' in the item is coded. Visuals can be desirable -- i.e. neutral, balanced or positive; or they can be undesirable --i.e. leaves the reader less likely to support, work for, or do business with the organization.

  ▷ "0=No visual

  ▷ 1=Desirable visual
     2=Undesirable Visual"

Key Messages

► Items were analyzed to determine if they contained one or more of the key messages. Key messages were provided by the clients

Message Tone

▶ Corporate Messages - Clear and apparent editorial positioning of a company favorably or unfavorably on its key messages. Each organization has its own key messages.

▶ Presence of a key message will not necessarily be explicit, but will convey the gist of the key message.

▶ The opposite of a key message is a negative message and should be coded as such. Examples were provided

▶ Negative conveys the opposite of the message

▶ Positive conveys the true meaning of the key message communicated

Message Integrity

▶ Message integrity was only coded in positive messages. Coding social media posts, also involves looking at pictures in the post to determine sentiment and messaging.

▶ There are four levels of integrity:

1. No Message/Negative message
2. Full = Item contains the complete message in its entirety
   ➲ Southwest has the best rates
   ➲ USWFS helps protect endangered animals
   ➲ Goodwill supports families

3. Amplified = Multiple messages are conveyed or inferred, all with positive tone and with potential amplifying adjectives.
   ➲ Awesome prices
   ➲ Best finds
   ➲ Best fishing advice
4. Partial = Only part of the message is inferred or communicated
   ➲ My daughter had a great time on our flight with the Southwest flight attendant
   ➲ USWFS was giving fishing lessons at the park today
   ➲ dropped my clothes off at Goodwill

Advocacy

▶ Each item was analyzed to determine if it advocated for the organization.

▶ Advocacy was either present or absen regardless if it was e negative or positive. Advocacy includes a call to action, or specific recommendation to act or change behavior or belief on the part of the reader (not the organization itself.)

  ⬧ e.g. "#iHateSouthwest - if u have had similar bad, inconsistent experiences with @SouthwestAir tweet them, they can't skate with their marketing"

  ⬧ e.g really SHOULD donate to Goodwill, fly Southwest or visit this park.

▶

▶

Intercoder Reliability Test

▶ Intercoder reliability refers to the level of agreement among coders when coding a corpus of messages using the same coding instructions and or codebook (Wimmer & Dominick, 2014).

▶ Krippendorff alpha was calculated because it can be used regardless of the number of observers, levels of measurement, sample sizes, and presence or absence of missing data.

▶ Krippendorff (2004) recommends an alpha level of at least .80 as a standard, accepting data in situations where.800 > α ≥ .667 —where tentative conclusions are still acceptable, and rejecting data where .667 ≥ α.

Krippendorf Alpha Results

Organization	Sentiment	Visual	Message	Message Tone	Integrity	Advocacy
Southwest	.745	.881	.583	.712	.656	.266
US Wildlife	.172	.034	.128	.096	.078	.051
Goodwill	.309	.07	.062	.028	.119	.146

Recommendations & Discussion

- Theoretically, if standards are valid and reliable, there should be no need to confer. However, accuracy and consistency increases when coders are allowed to to confer and it is common practice in many content analysis situations.

- Tailor messages to the platform or medium

  Messaging and advocacy occur very differently in Twitter than they did in Facebook. It may not be realistic to code for a single message across all platforms.

- Sentiment is more consistent and easier to code for consumer brands than for non-profits and government agencies.

- Any coding book needs to be thoroughly tested before coding begins.

- Do we revise the standard?