

Essay 2: Building Mistrust from Inconsistent Verbal and Non-Verbal Signals

W241, Section 2: Field Experiments

Kanitha Mann

Research Question

In-person communication amongst the general population is a combination of what is spoken and what is conveyed through body language and facial expression. These non-verbal signals tend to be consistent with verbal signals, giving the audience an accurate sense of what the speaker means. However, when verbal messaging is inconsistent with the speaker's non-verbal signals, the audience may misunderstand or develop a mistrust of the speaker. Such a scenario appears often in work and social settings to varying outcomes. Some examples include:

1. Someone who may be speaking but is distracted by their environment (for example, looking at items in their surrounding other than the audience) can make the audience feel like the weight of the words isn't as important as it would have been would the speaker make eye contact
2. A manager trying to say they're not upset with you for a mistake, but their voice is shaky
3. A Vice President of Sales who is pitching a good product to an important prospect, but doesn't sound convincing enough to win the sale

Each of these examples lead to the question: **do speakers who present inconsistent verbal and non-verbal signals lead their audience to mistrust them?** Furthermore, **is it possible to determine the point at which competing signals cause the audience to transition from simple misunderstanding to mistrust of the speaker?** Conducting an experiment pertaining to this question can grant us insight on what signals people rely on to decipher the "real" message the speaker meant. Furthermore, this exercise can help educate people on how to better conduct themselves in work and social situations so that they can communicate more effectively. Improved communication would lead to improved relationships by building trust and being transparent.

Subjects and Recruitment

The subjects for this experiment can either be recruited from [Amazon Mechanical Turk](#) (MTurk), where they will be promised a small compensation for their participation, or recruited to take a survey for a chance to win a gift card after completion. The best option will be the one where the most participants can be obtained within the constraints of the experiment budget. For example, subjects recruited through MTurk may only be promised \$0.50 for their time and feel like it is not worth the time. On the other hand, the probability of winning a \$50 gift card decreases as

more subjects are recruited, so potential subjects may feel like the low probability of winning is similarly not worth their time.

Randomization will occur depending on how subjects are recruited for the study. If conducted through MTurk, the subject will randomly be served with either the treatment or control versions of the experiment. However, if subjects are to be recruited to complete a survey, such survey vendors may not feature the capability to dynamically and randomly serve different versions through custom URLs. Different surveys for the control and the treatments will have to be created, and then distributed in separate ways. For example, if subjects were to be recruited from the I School's Slack community, one survey's URL would be placed in one channel, another survey's URL would be placed in another channel, and so forth. This method does not guarantee complete randomization, however, as subjects may have chosen which channels they choose to receive updates for. These methods will be the primary way in which the experiment will gain independence between subjects.

The Experiment

The subjects who have chosen to participate will be asked to watch a short (1-2 minute duration) videos of a person speaking. This speaker will be unobstructed so that the viewers may also watch their facial expression and body language. At the end of each video, subjects will be asked to rank the following questions on a Likert scale:

1. On a scale from 1 (least) to 5 (most), how well did you understand the speaker?
2. On a scale from 1 (least) to 5 (most), how well do you feel you trust the speaker?

The subjects will then be shown another video of the same speaker talking about something different and asked the same questions, for a total of ten videos.

The control group will be given ten videos in which the speaker's verbal and non-verbal signals are consistent. This set of videos will be known as Type A. This allows the experiment to set a baseline on how the audience judges the speaker when there are no competing signals to contend with.

The treatment changes these ten videos by introducing more examples of misaligned verbal and non-verbal signals. The treatments can vary in the following ways:

1. Type B: The first video shows consistent signals, followed by a video with inconsistent signals. This continues to alternate for the rest of the treatment. We would like to measure how subjects change their ratings when the speaker keeps alternating "tactics".
2. Type C: The first five videos shows consistent signals, while the second five videos shows inconsistent signals. We would like to measure if subjects change their ratings drastically after being exposed to the inconsistent videos at the end of the treatment.

3. Type D: The first five videos show inconsistent signals, while the second five videos shows consistent signals. We would like to measure if subjects can building understanding and trust after having seen the inconsistent videos at the beginning of the treatment.

We would ideally like to find enough videos so that all the subjects will view the same speaker. This can be found by taking a longer video and splicing it up to fill up examples for each type of video. The researchers will need to take great care in bucketing the video clips into their “correct” type. We also would like a speaker that may not be as well known, speaking about a subject that is not inflammatory. This avoids subjects already having a well cemented opinion of more colorful characters in our present world, and also avoiding topics that are divisive.

Table 1. Experiment With Treatment Varying in Type of Contrary Videos

	$Y_i(0)$	$Y_i(1)$		
		Treatment 1	Treatment 2	Treatment 3
N subjects	<i>Type A</i>	<i>Type B</i>	<i>Type C</i>	<i>Type D</i>

Type A: 10 videos in which the speaker has aligned verbal and non-verbal signals

Type B: Alternating of 10 videos showing both aligned and misaligned verbal and non-verbal signals

Type C: 5 videos of aligned signals, followed by 5 videos of misaligned signals

Type D: 5 videos of misaligned signals, followed by 5 videos of aligned signals

One potential covariate that may be useful to explore is to ask subjects what their work experience is. Subjects who are in higher roles in white collar companies or who have years of experience in a customer service role may be capable of discerning a speaker’s true message, in spite of inconsistent verbal and non-verbal signals. On the other hand, subjects who are just starting a white collar position or who do not work in customer service may not have this experience. Collecting this data may be used to dynamically place subjects into blocked or clustered groups.

Outcome Measures

The primary outcome measure will be the mean value for the two questions posed to each treatment group. The Likert scale will also provide additional measures, including the distribution of the ratings, and the ratings per video as it sequentially progresses throughout the treatment. There are many people who maintain that using parametric statistical methods on Likert scale data are meaningless, due to the ordinal nature of the scale. For example, one subject’s “1” rating about understanding may be the same as another subject’s “2” rating about understanding, so using the numbers 1 and 2 to weigh the responses is disingenuous.

However, there are resources¹ available to help researchers effectively use statistical methods on Likert scale data.

Risks

How a viewer reacts to a speaker's message - consistent signals or not - is still a subjective topic. Aside from more severe limitations like spectrum disorders, blindness, and deafness, there are certain factors that will influence an individual's decision making skills when deciding whether a speaker is understanding and truthful. This includes their upbringing, cultural norms, and gender norms. Researchers may wish to collect information on subjects, like gender and age, in order to account some of these biases in the analysis, along with collecting information on work experience.

The types of videos chosen for this analysis must also be consistent for the experiment. Speakers that employ sarcasm are mindful of the inconsistencies between speech and body language, so these examples of intentional misalignment should not be used. As mentioned earlier, videos that feature particularly divisive speakers or subjects should not be used, as subjects' opinions may be well-cemented long before the experiment is conducted.

¹ "Likert scales, levels of measurement and the "laws" of statistics"
<https://link.springer.com/article/10.1007/s10459-010-9222-y>. Accessed 2 Oct. 2018.