Supplementary Materials B:

Codebook for identifying misunderstandings in online dialogue (Version 1)

Author note:

This codebook accompanies the article titled: *The Repeated Adjustment of Measurement Protocols (RAMP) method for developing high-validity text classifiers.* It is the first version of the codebook used in the RAMP method and, consequently, used for the non-iterative method.

Introduction

This coding manual outlines instructions for identifying sentences containing explicit verbal expressions of misunderstanding. The coding framework is adapted from the interpersonal perception method (Laing et al., 1966). The method distinguishes between *direct perspectives* (I think X), *meta-perspectives* (I think you think X) and *meta-meta-perspectives* (I think you think I think X). Table 1 demonstrates how these different levels of perspective taking can be combined to determine various social-psychological variables: agreement, direct understanding, and felt understanding.

Table 1: Perspective-taking as the basis for agreement, understanding, and misunderstanding

Phenomenon	**Self**	**Other**	**Matching**
Agreement	**Direct Perspective** I think X	**Direct Perspective** I think X	Yes
Disagreement	**Direct Perspective** I think <u>not X</u>	**Direct Perspective** I think X	No
(Direct) understanding	**Meta-Perspective** I think you think X	**Direct Perspective** I think X	Yes
(Direct) misunderstanding	**Meta- Perspective** I think you think <u>not X</u>	**Direct Perspective** I think X	No
Felt understanding	**Meta-Meta Perspective** I think you think I think X	**Meta-Perspective** I think you think X	Yes
Felt misunderstanding	**Meta-Meta Perspective** I think you think I think <u>not X</u>	**Meta-Perspective** I think you think X	No

This coding scheme is concerned with identifying instances of direct and felt misunderstanding in naturally occurring online dialogue data.

Coding scheme

Datasets

This coding manual was designed and tested using the following three datasets:

I. **Reddit**: random sample of 1,000 posts + comments from 30 subreddits gathered through API. Each subreddit has different rules and norms that participants abide by. There is a wide variation of topics and types of discussion on reddit. See example here: https://www.reddit.com/r/changemyview/comments/s4opx2/cmv_i_dont_see_how_a_person_can_be_a_christian/

II. **Wikipedia talk pages**: pre-existing dataset (Danescu-Niculescu-Mizil et al., 2012) involves Wikipedia editors discussing articles and edits. Often these take place on an editors own talk page, meaning that people direct questions or comments to them. See example here: https://en.wikipedia.org/wiki/User_talk:U3964057 [NOTE: the format is strange, a new turn is indicated by an indent and ends in an editor signing off their name].

III. **Twitter**: pre-existing dataset (Thought Vector & Axelbrooke, 2017)involves dialogues between organisations & customers. See example here: https://twitter.com/SueHiles/status/1486337775746093062

Unit of analysis

The unit-of-analysis for coding misunderstandings is sentences, sampled from a dialogue "turn" (i.e., a comment or post). Misunderstanding is coded as a binary variable, where "1" refers to the presence of misunderstanding. These are coded in an excel spreadsheet with the structure of Table 2.

Table 2: Example code structure

turnId	Sentence	Misunderstanding
t0	Example sentence 1	**0**
t295	Example sentence 2	**1**
t365	Example sentence 3	**0**

Table 2 Example code structure

Coding is performed blind to context. This means that sentences are isolated from the context of its parent turn or previous turns by other participants.

Coding Misunderstanding

Misunderstanding is defined as the presence of *direct* or *felt* misunderstanding in a sentence. Direct misunderstanding is defined as a participant indicating confusion about a direct perspective (meta-perspective) given by a participant in a previous turn. Felt misunderstanding is defined as a participant acknowledging that the reply from a previous participant has misunderstood a previous turn of theirs (meta-meta-perspective). A misunderstanding should be identifiable in a sentence as addressing *another person's perspective*.

Do code for misunderstanding when:

- a participant makes an overt statement showing misunderstanding of a previous participant's perspective (e.g. "I don't get you're point", "I don't get it")
- a participant overtly states that another participant in dialogue has misunderstood them (e.g. "you've misunderstood me")
- a participant asks for overt clarification (e.g. "what did you mean when you said", "did you mean")
- a participant makes an explicit third turn repair (e.g. "I actually meant", "what I meant was", "I didn't mean that")

Do *not* code for misunderstanding when:

- a participant is asking for clarification about a topic (e.g. "do you know what X means?")
- a participant is asking for elaboration about a topic (e.g. "can you tell me more about X?")
- a participant is asking for new information / new behavior from the other (e.g. "can you DM me your details", "on a different note, do you know about X?").
- The misunderstanding is about information rather than an other's opinion / view of that information.

Examples of direct misunderstanding:

I misunderstood you, I don't understand what you mean, I don't understand, no idea what you're saying, what did you mean, that doesn't make sense, that makes no sense, we're misunderstanding each other, can you clarify what you meant

Examples of felt misunderstanding:

I didn't mean, you misunderstood me, I wasn't saying, I wasn't saying that, I never meant to, you missed my point, that's not what I meant, you took this the wrong way, that's not what I was talking about, I wasn't talking about, my comment was meant to

NOTE: also consider permutations of these phrases with adverbs, e.g. *you've completely misunderstood me, that really doesn't make sense*

References

Danescu-Niculescu-Mizil, C., Lee, L., Pang, B., & Kleinberg, J. (2012). Echoes of power: Language effects and power differences in social interaction. *Proceedings of the 21st International Conference on World Wide Web*, 699–708. https://doi.org/10.1145/2187836.2187931

Laing, R. D., Phillipson, H., & Lee, A. R. (1966). *Interpersonal Perception: A Theory and a Method of Research* (pp. vii, 179). Springer.

Thought Vector, & Axelbrooke, S. (2017). *Customer support on Twitter (v10)*. https://kaggle.com/thoughtvector/customer-support-on-twitter