

Explanatory Style Annotation Guide – DRAFT

HiLT Lab

March 6, 2018
v0.8.2

Contents

1	Event Attribution Units (EAUs)	2
1.1	Events	2
1.2	Attributions	3
2	Annotating Event-Attribution Units and their Features	3
2.1	Annotation Phases	3
2.1.1	Phase 1: Event spans	3
2.1.2	Phase 1B: Event polarity	3
2.1.3	Phase 2: Attribution spans	4
2.1.4	Phase 3: Attribution dimensionality	4
2.2	Annotation span conditions	5
2.2.1	Events	5
2.2.2	Attributions	6
2.2.3	All annotations	6
2.2.4	EAUs	6
2.3	Annotation tips	7
2.3.1	EAUs	7
2.3.2	Attribution Dimensions	8

Research suggests that the way that people explain things that happen in their lives can be revealing.¹ A consistently pessimistic *explanatory style*, for example, can help indicate depression. The ability to detect these explanations, and consequently describe a person's explanatory style, is therefore a valuable skill for therapy-minded Companionbots since it lets them know when and about what to start therapeutic dialogue.

1 Event Attribution Units (EAUs)

In order to describe a person's explanatory style, we need a sample of how that person explains things that happen to them, particularly things about which the person has a non-neutral opinion. This sample will be made up of a set of *event-attribution units*² (EAUs) and some descriptive features about them. EAUs are the conjunction of an *event* and an *attribution* which explains the existence of that event.

1.1 Events

Schulman, Castellon, and Seligman provide a definition of *event* in their annotation guidelines³:

An event is defined as any stimulus that occurs in an individual's environment or within that individual (e.g. thoughts or feelings) that has a good or bad effect from the individual's point of view. Events can be mental (e.g. I was afraid), social (e.g. I got a pay raise) or physical (e.g. I got in a car accident). Events should be unambiguously good or bad from the individual's point of view and may occur in the past, present or hypothetical future. Events that have good and bad elements, neutral events or events that do not affect the S [the individual] should not be extracted.

While we do attempt to follow this definition as closely as possible, we find that it is easier to reproduce similarly formed *events* from our own transcripts by adding several conditions to our annotations (§2.2), supplementing those outlined in Schulman et al's guidelines.

¹Schulman, P., Castellon, C., & Seligman, M. (1989). Assessing explanatory style: The content analysis of verbatim explanations and the attributional style questionnaire. *Behavior Research and Therapy*, 27(5), 505-512

²appendix of CAVE paper, with guidelines

³see footnote 2

1.2 Attributions

Attributions are explanations that people give for the good and bad events that happen in their lives. These attributions are causal in nature, such that the attribution is described in a way that suggests it caused the event to happen.

For the purposes of our annotation, it is essential 1) that attributions be proposed by the subject who we are interested in (i.e. the one whose subjectivity we evaluate when deciding whether to classify a given event as either good or bad) and 2) that the causal relationship identified was clearly intended to be communicated as a feature of conversation.

2 Annotating Event-Attribution Units and their Features

2.1 Annotation Phases

2.1.1 Phase 1: Event spans

Individual Annotation

1. At least two annotators annotate the most appropriate text span of each event that they identify to be part of an EAU, labeling each as type Event. Annotators will be provided an annotation document containing pre-made heuristic annotations. Only spans which overlap with any part of a sentence which contains a given heuristic annotation are eligible for annotation during this phase.

Consensus Annotation

1. Annotators discuss each given annotation and copy to the consensus set those which they agree meet all given annotation conditions, including the existence of an associated attribution.
 - (a) Annotators may decide to include any appropriate annotation (whether it be made during individual annotation or not) in the consensus set that they may happen to notice during discussion.

2.1.2 Phase 1B: Event polarity

Individual Annotation

1. One annotator annotates the polarity of each Event given in the consensus set from Phase 1 as either Positive or Negative as a feature of the annotation.

2.1.3 Phase 2: Attribution spans

Individual Annotation

1. For every Event annotation from Phase 1B, two annotators individually annotate the text span of every associated attribution as type Attribution.
 - (a) Provide the ID number of the associated event as a feature, Caused_Event.

Consensus Annotation

1. Annotators discuss each given annotation and copy to the consensus set those for which they agree meet all given annotation conditions.
 - (a) Annotators may decide to include any appropriate annotation (whether it be made during individual annotation or not) in the consensus set that they may happen to notice during discussion.

2.1.4 Phase 3: Attribution dimensionality

Individual Annotation

1. For every Attribution within the given consensus set, at least two annotators annotate values for each of the three dimensions, Personal--External, Permanent--Temporary, Pervasive--Specific, as features, using a digit between 1 and 7.

Consensus Annotation

1. Annotators discuss each given annotation and copy to the consensus set those which they agree meet all given annotation conditions.
 - (a) Annotators may decide to include any appropriate annotation (whether it be made during individual annotation or not) in the consensus set that they may happen to notice during discussion.

2.2 Annotation span conditions

Some definitions:

propose, *v.* to make a statement about something which could be evaluated as either *true* or *false*, e.g. *a bell rang repeatedly*, but not just *a bell* or *to ring* or *repeatedly*.

corefer, *v.* to share a single reference, e.g. *Paul* and *He* in “*Paul* had a great time” and “*He’s* always in a good mood.”

backchannel, *v.* to affirm active listening, e.g. “okay,” “uh-huh,” “I’m listening,” etc.

turn *n.* in conversation, an uninterrupted span of speech by a single person (excepting any *backchannel*).

2.2.1 Events

For every **event** annotation:

1. The *event* must have been mentioned during the given conversation.
2. The event must have been or otherwise be very probable to have been evaluated as either good or bad by subject in terms of its effects on the subject.
3. The text span that expresses the *event* must be in the form of a complete sentence or otherwise unambiguously *corefer* with a complete sentence present in the transcript expressing such event when no other non-pronoun mention of the event appears within the same *turn* as the otherwise most appropriate event mention.
4. The text span that expresses the *event* must primarily *propose* that event.
5. Experience of the *event* by the participant must not rely solely on a generic statement, e.g. not “Teachers often feel stressed” even if the individual is in fact a teacher.
6. The span must overlap with a supplied heuristic annotation.
7. The text span expressing the *event* must be the minimal length necessary to satisfy all preceding conditions.

2.2.2 Attributions

For every *attribution* annotation:

1. The *attribution* must have been mentioned explicitly during conversation.
2. The text span that expresses the *attribution* must primarily suggest that attribution.
3. The causal relationship identified between the attribution and its associated event must have clearly been intended to have been communicated.
4. The internality dimension of the attribution must be unambiguously agentive, i.e. the attribution must be caused at least in part by some entity.
5. *Attribution* annotations must be the minimal length necessary to satisfy all preceding conditions.

2.2.3 All annotations

For every annotation made:

1. Annotations may be non-contiguous.
 - (a) Use *annotation_continuation*, replacing *annotation* with the type name of the continued annotation, e.g. *Event_continuation*.
 - (b) Annotation continuations must not be interrupted by a complete annotation of the same type.
2. Annotation spans may fail to provide all relevant information and remain valid only if nearby contextual information relieves the insufficiency.
3. Avoid creating an annotation such that sentence-ending punctuation falls within and at the end of the annotation.

2.2.4 EAUs

Every *EAU*:

1. must consist of only the nearest pair of Event and Annotation mentions which each satisfy their respective conditions, except:

- (a) when an attribution is repeated in such a way that does not seem to merely be a repair (e.g. stuttering, clarification, etc.). In such cases, annotate the Event using one span, and annotate each Attribution with separate spans, each referring back to the same Event span.
- 2. must not overlap in its event and attribution spans.
- 3. given a contiguous series of similar events which an attribution or series of attributions with similar internality explains, annotate a single event attribution, grouping like events into a single event span and like attributions into a single attribution span.

2.3 Annotation tips

2.3.1 EAUs

Daisy-chaining It is acceptable for the same span of text to be selected as the caused Event in one EAU and separately as an Attribution as part of a different EAU.

Example:

E: I just had a thirty minute panic attack.

A: My credit card was not where I usually keep it.

E: My credit card was not where I usually keep it.

A: I had changed the location.

Concept repetition It is acceptable to annotate the same conceptual EAU as many times as it is mentioned, but not when it seems only to have been repeated for the sake of conversational clarity.

Conversational politeness Sometimes people say things only out of a sense of politeness, e.g. apologizing for a minor disruption in conversation. If you get the sense that a remark was only made out of an effort to save or maintain face, or does not otherwise represent a sincerely felt effect, do not consider it an Event.

Example:

E: I'm sorry

A: I couldn't answer your call.

NB: This proscription does not preclude the possibility that some polite speech acts might indicate the sincerely felt effects of an Event.

2.3.2 Attribution Dimensions

1. Internality

Who is to blame? Who takes the credit? In answering the question, “What degree of internality/externality is present in this Attribution?” you will necessarily be answering the question, “According to the speaker, who/what is to blame (in the case of a negative event) or who/what should take credit (in the case of a positive event) for the associated Event?”

If there seems to be more than just the speaker responsible for a given event, try to identify who has more focus, and adjust annotation accordingly. A heuristic for identifying who has more focus is identifying which of those responsible are mentioned first.

Example:

E: The house looks marvelous.

A: I had the cleaning ladies over.

Internality:5/7; “I” has more focus than “the cleaning ladies.”

Example:

E: The house looks marvelous.

A: The cleaning ladies came over.

Internality:3/7; Though the cleaning ladies are the only ones explicitly mentioned, it is understood within the context that the speaker must have taken some action for them to come over at all, therefore precluding annotation from being *highly* external.

Example:

E: I can’t keep myself motivated

A: because of my depression.

Internality:7/7; The speaker refers to an illness of theirs by means of a personally referring pronoun, which is different from...

Example:

E: I can’t keep myself motivated

A: because of the depression.

Internality:5/7; Though the depression is known to belong to the speaker, the speaker chooses not to use any personal reference when mentioning the depression.

Example:

E: You can't keep yourself motivated

A: because of your depression.

Internality:5/7; Assuming that the context suggests that the information relevant to this generic second person mentioned (i.e. the "(you)'s") is actually highly relevant to the speaker, the speaker chooses not to use any properly personal reference when mentioning the depression.

2. Stability

Temporally constrained Attributions, temporally abstracted Events

When annotating the stability of an Attribution, you are essentially annotating the likelihood that the Attribution will be considered a cause of the Event by the speaker if the Event were to happen again.

In answering this question, it will be necessary to hypothesize about future occurrences of the Event, but be careful not to extend this hypothetical way of thinking to the Attribution itself. That is to say that in answering this question, your mental model should consider the Attribution constrained to the particular instance mentioned in time if there is any temporal reference, but consideration of the Event should be abstracted to all possible future occurrences of the Event, regardless of how unlikely it really would be to happen again.

Example:

E: I can't focus in class

A: because of my ADHD.

Stability:7/7; Even if the speaker is not likely to remain a student for much longer, which would make it very unlikely that this Event would ever happen again, the Attribution must be analyzed within the context of the Event *actually happening*.

Consider phrasing When annotating attribution dimensions, consider the way the person decided to phrase the attribution.

Example:

E: I can't keep up at Bingo

A: because of how old I am.

Stability:7/7

Example:

E: I can't keep up at Bingo

A: because I'm ninety-one years old.

Stability:4/7

3. Globality

Avoid considering butterfly effects Annotating Globality essentially is answering the question, "How wide is the scope of things believed to be affected by this Attribution?" Inevitably, some Attributions have little prior dialogue to rely on in answering this, in which case Schulman et al's guidelines prompt you to mentally model an average person as a surrogate upon which to gauge the Attribution's effects.

When doing this, try to only count the obvious immediately following effects and not all of what might likely be chain reactions of each of those effects. For example, if given an Attribution to the effect of "They're changing my blood pressure medication," the only obvious domain in which that plays a major, immediate role is the speaker's physical health, though a wide range of future changes might possibly or probably change as a result. However, the majority of these further effects would likely only be a result of whatever physical change the medication brings about, and not the change in medication itself, therefore the Attribution ought to be considered more Specific than Global.

All this isn't to say that an Attribution with an otherwise narrow scope of effect could never be annotated as more pervasive; if you believe that the speaker believes it's pervasive, mark it as such.