

From Categories to Decisions: A Framework for Attitudinal Analysis of Evaluative Language

Jiamei Zeng^{1*}, Haitao Wang^{2*}, Harry Bunt³, Xinyu Cao², Sylviane Cardey⁴, Min Dong⁵, Tianyong Hao⁶, Yangli Jia⁷, Kiyong Lee⁸, Shengqing Liao⁹, James Pustejovsky¹⁰, François Claude Rey⁴, Laurent Romary¹¹, Jianfang Zong², and Alex C. Fang^{1**}

¹ City University of Hong Kong, PR China (jiameizeng3-c@my.cityu.edu.hk, alex.fang@cityu.edu.hk)

² China National Institute for Standardization, PR China ({wanght, caoxy, zongjf}@cnis.edu.cn)

³ University of Tilburg, The Netherlands (Harry.Bunt@tilburguniversity.edu)

⁴ University of Franche-Comte, France ({francois_claude.rey, sylviane.cardey}@univ-fcomte.fr)

⁵ Beihang University, PR China (mdong@buaa.edu.cn)

⁶ South China Normal University, PR China (haoty@m.scnu.edu.cn)

⁷ Liaocheng University, PR China (jiayangli@lcu.edu.cn)

⁸ Korea University, Korea (ikiyong@gmail.com)

⁹ Fudan University, PR China (sqliao@fudan.edu.cn)

¹⁰ Brandeis University, USA (jamesp@cs.brandeis.edu)

¹¹ National Institute for Research in Digital Science and Technology, France (laurent.romary@inria.fr)

Abstract

The study reported in this paper aims to contribute to the development of an annotation scheme for evaluative language, based on Appraisal Theory, that addresses key sources of classification problems. In particular, it aims to develop a unified annotation scheme that proposes (1) a three-component annotation model comprising Appraiser, Appraised and Appraisal Element, (2) the operationalised distinction between Affect and Appreciation governed by a criterion of experienter salience and a criterion distinguishing personal emotions from evaluations of conduct, and (3) a decision framework for the Judgement-Appreciation distinction structured on the target and lexis types operating through override conditions and substitution tests. The revised framework is illustrated with examples selected from a corpus of news discourse in English and is designed to be replicable across future Appraisal-based studies of evaluative language.

Keywords: evaluative language, appraisal theory, affect, judgement, appreciation, corpus

1. Introduction

Appraisal Theory (Martin & White, 2005) provides one of the most widely adopted frameworks for analysing evaluative language in discourse, and its Attitude system, comprising Affect, Judgement, and Appreciation, has been applied extensively in corpus-based studies of media, political, and academic discourse (Fuoli, 2018; Read & Carroll, 2012). Despite this broad uptake, the annotation of Attitude in practice remains a source of persistent difficulty. In particular, three problems have proven resistant to resolutions within existing annotation methodologies. First, most annotation schemes mark up only the evaluative expression itself without formally specifying its source (the Appraiser) and target (the Appraised), leaving the relational structure of evaluation implicit and creating conditions for undetected disagreement (Read & Carroll, 2012; Fuoli, 2018). Secondly, while theoretically motivated by the concept of ‘institutionalised Affect’ (Martin, 2000; Martin & White, 2005), the boundary between Affect and the

other two subsystems lacks operationally explicit criteria for deciding when an expression has crossed from felt emotions into Judgement or Appreciation. Thirdly, the boundary between Judgement and Appreciation, particularly for abstract human-derived entities such as policies, plans and reports, depends on the relationship between the evaluative lexis and its target in ways that existing accounts describe but do not fully proceduralise (Bednarek, 2009; Martin & White, 2005; White, n.d.).

The present study seeks to address these three problems by developing a unified annotation framework grounded in Appraisal Theory. In an earlier companion study (Zeng et al., 2025), we reported the results of an annotation experiment applying this framework, including inter-annotator agreement measures and quantitative evaluation. The present article addresses the annotation principles and decision procedures for distinguishing among the three major Attitude categories at the category level. In what follows, the framework is described and illustrated with

* Equal contribution

** Corresponding author

examples drawn from a corpus of English-language news discourse on the COVID-19 pandemic. The aim is not to report corpus findings but to demonstrate the annotation procedures themselves, providing a practical and replicable framework for attitudinal analysis in the study of evaluative language.

2. The Annotation Framework

2.1 Why annotate Appraiser, Appraised, and Appraisal Element together

Most existing corpus annotation work in the Appraisal tradition centres on the identification and classification of evaluative expressions, what we term the Appraisal Element (AE), without formally indicating or marking up who is performing the evaluating or the target of the evaluation. This reduced approach has been the practical default in large-scale annotation studies (Fuoli, 2018; Read & Carroll, 2012). However, this reduction has caused theoretical and practical problems. Theoretically, evaluative meaning is inherently relational: the classification of an evaluative expression depends on the source (who is evaluating) and the target (what/who is being evaluated), not on the lexical item alone (Martin & White, 2005). Practically, omitting the source and target of evaluation makes it impossible to recover who is evaluating whom in multi-voice discourse (White, 2012).

The present study addresses these interconnected problems through a proposal to identify three relational components of appraisal A: AS (appraisal source), AT (appraisal target), and AE (appraisal element, i.e. the linguistic realisation), thus $A = \{AS, AT, AE\}$. The idea of identifying AS and AT alongside AE is not new in itself. It is already implicit in the theoretical literature (Martin & White, 2005). What has been lacking is a schematic requirement that all three components be considered and marked up where possible for every evaluative instance, together with a principled procedure for resolving the classificatory ambiguities that arise when multiple readings become available. In the rest of this article, each example of evaluation is annotated with triples in the form of $\langle AS \rangle$, $\langle AE \rangle$ and $\langle AT \rangle$. Unless otherwise indicated, all the examples in the article are selected from a corpus of COVID-19 news reports (Zeng et al, 2025). The source of each example is provided in round brackets and coded in the format year-month-area-media.¹

2.2 What each component captures

The Appraiser is the semantic source of the evaluative act (Martin & White, 2005, pp. 71-81). It

can be explicitly realised, as when a named individual or entity is grammatically present, or implicitly realised, as when the evaluation is attributable to the authorial voice. In the Affect subsystem, AS corresponds to the experiencer of emotion. AT is the semantic target of evaluation. It can be a person, behaviour, thing, process, or proposition. In the Affect subsystem, AT functions as the stimulus triggering the emotional response. AE is the linguistic realisation of evaluation, articulated by AS about AT. Its classification carries information about the attitudinal categories including Affect, Judgement, and Appreciation, and their subcategories. Their relational structure can be represented as $\langle AS \rangle$ - $\langle AE \rangle$ - $\langle AT \rangle$, that is, AS evaluates AT as possessing a particular attitudinal quality linguistically encoded through AE.

Consider [1] and [2], which illustrate how the components work together and how the same formal structure applies across attitudinal categories.

[1] "I think it was just a really bad idea," Ms. Ratley said. (20-08-US-NT)

$\langle AS \rangle$: Ms Ratley
 $\langle AE \rangle$: Appreciation: Reaction: *really bad*
 $\langle AT \rangle$: *idea*

AS is identified with Ms. Ratley, to whom the evaluative proposition is attributed through the reporting clause. The attribution frame simultaneously carries Engagement significance but the present study focuses exclusively on the Attitude layer. AT is *the idea*, a non-human entity, and AE *really bad* realises negative Appreciation in the Reaction subcategory. The evaluative triple can be represented as a tree structure in Figure 1, where the linguistic encodings of evaluation are anchored in terminal nodes and annotated by the non-terminal nodes for appraisal types and categories.

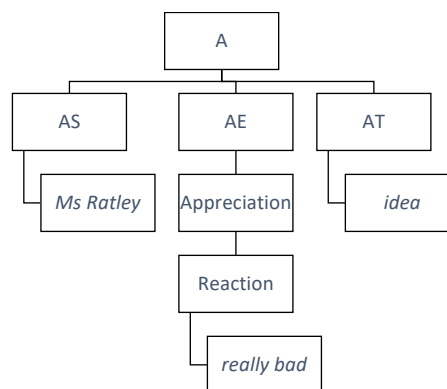


Figure 1: A tree representation of the evaluative structure of Example [1]

¹ The area codes include CN (China), HK (Hong Kong), SG (Singapore), UK and US. Media codes include CD (China Daily), NT (New York Times), SM

(South China Morning Post), ST (Strait Times), and TG (The Guardian).

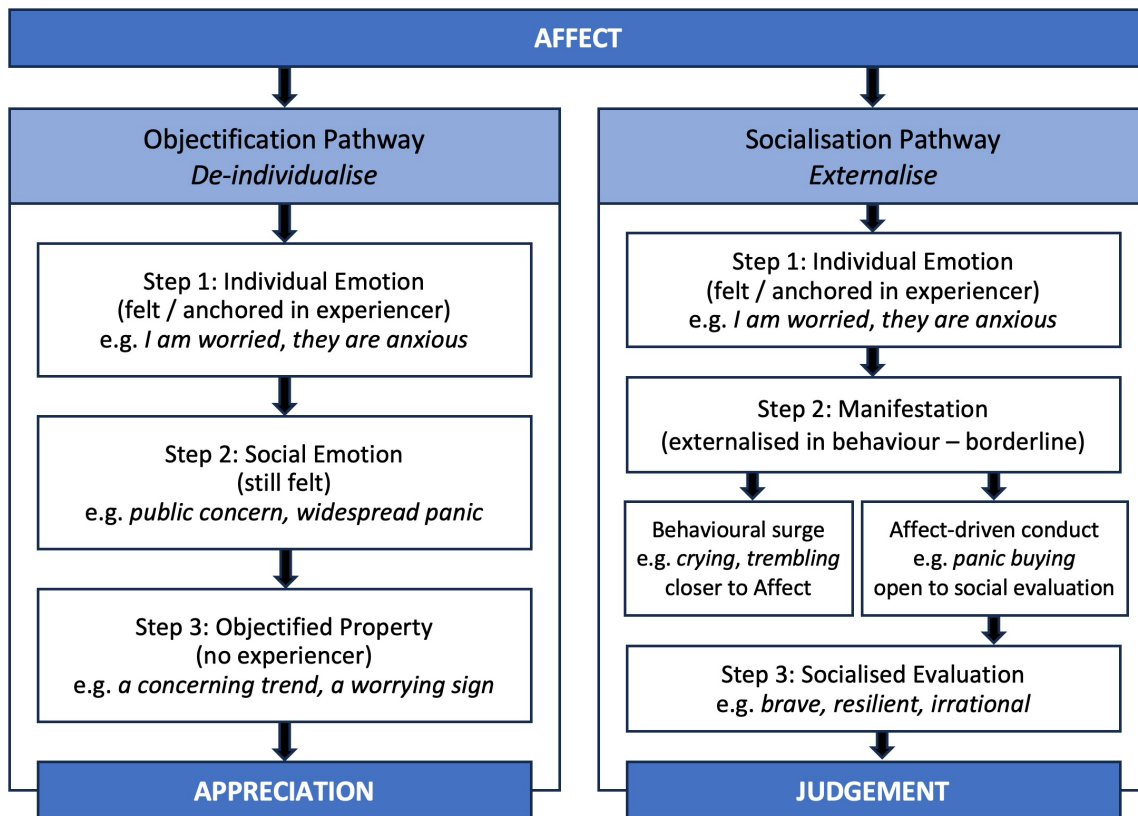


Figure 2: Pathways of Affect towards Appreciation and Judgement

Differently, AE in [2] below indicates an attitudinal evaluation in the Affect sub-system.

[2] “They fear side effects. They don’t trust vaccines in principle, or they want to wait and see what happens to other people first,” said Denis Volkov, deputy director of the Levada Center. (21-06-US-NT)

(AS: *they*)

(AE: Affect: In/security: *fear*)

(AT: *side effects*)

AS *they* is the experiencer, AT *side effects* is the trigger, the stimulus triggering the emotional response, and AE *fear* realises negative Affect in the In/security subcategory.

3. Distinguishing Affect from Judgement and Appreciation

3.1 Theoretical foundation: Two pathways from Affect

The three Attitude subsystems, Affect, Judgement, and Appreciation, share a common evaluative root rather than occupying wholly separate domains. As Martin (2000) argues, ‘AFFECT can perhaps be taken as the basic system, which is then institutionalized in two major realms of uncommon sense discourse’: Judgement represents Affect recontextualised as an evaluation matrix for behaviour while Appreciation represents Affect

recontextualised as an evaluation matrix for the products of behaviour and the natural world. White (2004) clarifies the practical distinction: evaluative language can be divided into emotion, in which attitudinal assessments are indicated through descriptions of emotional reactions or states of human subjects, and opinion, under which a positive or negative quality is said to be an inherent property of the phenomenon being evaluated. Martin and White (2005) formalise this relationship as institutionalised Affect.

This insight explains why the boundaries between subsystems are not clear-cut: if Judgement and Appreciation are derived from Affect, expressions at the boundary will inevitably retain traces of both. However, ‘institutionalised Affect’ as a formulation remains too schematic to resolve the specific boundary decisions that arise in practical annotation. We therefore aim to refine this schema by describing the relationship between Affect and the other two subsystems as two directional pathways, each governed by a distinct mechanism. Both pathways begin from the same point of departure, in which feeling moves beyond the private experience of an individual. What distinguishes the two pathways is the direction of that movement. One pathway proceeds toward the objective world while the other toward the social world. Figure 2 illustrates the two pathways and their respective stages.

The first pathway from Affect to Appreciation can be described as individual emotion → social emotion → objectification. At the initial stage, Affect is anchored in a specific experiencer and the evaluative meaning remains at the level of private emotional feelings. In the next stage, emotions expand beyond individual experience into a collectively circulating condition: expressions such as public concern no longer refer to a single experiencer and yet their semantic core still lies in the existence and circulation of emotions within a social field. The emotion has become de-individualised but it has not yet been fully detached from human feelings. In the final stage, affective meaning undergoes objectification: what was originally felt by people is reconfigured as a quality of an entity, trend, situation, atmosphere, or event, as in a worrying sign. Discourse no longer foregrounds who is feeling worried but the object as possessing a worrying quality in itself. The emotional response has been encoded as an attribute of the evaluated entity.

The second pathway, from Affect to Judgement, can be described as individual emotion → individual manifestation → socialised evaluation. Unlike the movement toward Appreciation, which proceeds by objectification, the movement toward Judgement proceeds by socialisation. It begins with felt emotions, which may then become externalised in the form of individual manifestations. Such a manifestation can be divided into two subtypes. The first consists of pure manifestation, corresponding broadly to what is termed 'behavioural surge': crying, trembling, or freezing (Martin and White 2005, p. 47). These are bodily or behavioural manifestations of emotion that start to be publicly visible displays rather than remaining purely internal states. The second subtype consists in an affect-driven social conduct, that is, emotion-driven action that has an impact on others or on the social environment, such as expressions like *panic buying*. Unlike crying or trembling, such conduct is not only expressive but also socially impactive, triggering social or collective response, thus readily open to public evaluation. In the final stage, these manifestations are interpreted as signs of what is classified as social esteem or social sanction (Martin and White 2005, pp. 52-56), that is, as indicators of a person's normality, capacity, tenacity, veracity, or propriety, thereby entering the domain of Judgement. Judgement does not arise directly from emotion itself but from the social reading of emotional manifestations. What moves Affect toward Judgement is not the emotional feeling alone but the fact that feeling becomes publicly visible, thus interpretable and impactive according to social norms.

Following from here, we formulate two operational criteria in Sections 3.2 and 3.3 below.

3.2 Criterion 1: Felt Experience or Objectified Property (Affect vs. Appreciation)

The first annotation criterion follows directly from the objectification pathway, that is, an expression realises Affect when an identifiable experiencer, whether an individual or a collective human entity, is present in whose emotional state the evaluation is grounded. When the emotional quality has been fully detached from any experiencer and attributed to an entity, situation, or text as an inherent property, the expression realises Appreciation. This criterion operates across a gradient of experiencer salience, from individuals to collectivity to absence. The clearest cases are those in which the experiencer occupies subject position and AE directly predicates an emotive state. Consider

[3] Now, we are grieving, afraid and confused. (20-05-US-NT)

<AS: *we*>

<AE: Affect: Un/happiness: *grieving*>

<AT: unspecified>

<AS: *we*>

<AE: Affect: In/security: *afraid*>

<AT: unspecified>

<AS: *we*>

<AE: Affect: Dis/satisfaction: *confused*>

<AT: unspecified>

AS *we* is the explicit experiencer and *grieving*, *afraid* and *confused* each directly attribute an emotive state to it, yielding three co-occurring AEs belonging to the Affect category. No trigger is syntactically realised; this illustrates what Martin and White (2005, p. 47) call undirected mood, an emotional state reported without a specific stimulus. The absence of any identifiable AT is itself a diagnostic indicator: Judgement and Appreciation by definition require an entity being evaluated, and its absence strongly favours an Affect classification.

Between the implicit experiencer and full objectification lies a middle zone: social emotion, in which an emotional state is no longer anchored in a specific individual but still circulates as a felt condition within a collectivity.

[4] Criticism and worry continue to plague other colleges hoping to offer in-person learning this fall. (20-08-US-NT)

<AS: *other colleges*>

<AE: Affect: In/security: *worry*>

<AT: *in-person learning this fall*>

Worry is an emotion noun. While no individual experiencer is named, worry here is not presented as a property of the colleges or of the learning environment. It is presented as a social-emotional condition actively afflicting the colleges, as the verb plague makes clear. The experiencer is a

recoverable collective (university communities, administrators, faculty), making this a case of social emotion rather than fully objectified Appreciation. The absence of an explicit individual does not disqualify Affect; what matters is whether the emotional meaning remains grounded in someone experiencing it, even if that someone is a collectivity.

When the emotional quality has been fully detached from any experiencer and functions to characterise an entity or text, the expression realises Appreciation.

[5] This “nightmare scenario” may come about. (20-04-SG-ST)

(AS: implicit author)
(AE: Appreciation/Reaction: *nightmare*)
(AT: *scenario*)

AE *nightmare* is originally an Affect-laden term, but here it functions as a modifier attributing a quality to the scenario itself. This parallels the contrast that White (n.d.) draws between *the building bores me*, where the evaluative quality is anchored in a human experiencer, and *the building is boring*, where the emotion has been detached from any human experiencer and reattached to the entity as an intrinsic property. In [5], *nightmare* functions in precisely this second way: no specific or collective experiencer is construed, and the clause foregrounds the object (*scenario*) and its inherent affective charge rather than anyone’s experience of it.

The contrast between Examples [4] and [5] is instructive. In [4], *worry* is something that the colleges have or suffer; in [5], *nightmare* is a quality that the scenario is said to possess. The former foregrounds an experiential state and the latter foregrounds an entity property. This distinction marks the boundary at which Affect gives way to Appreciation along the objectification pathway.

3.3 Criterion 2: Personal Emotion or Evaluation of Conduct (Affect VS. Judgement)

The second annotation criterion follows from the socialisation pathway: an expression realises Affect when it foregrounds a personal emotional state, whether a mood, disposition, or surge, and realises Judgement when the primary evaluative content concerns how someone behaved, assessed against the dimensions of social esteem or social sanction. The boundary between the two is particularly subtle when an expression can be read as either an individual’s feeling or an assessment of conduct, a difficulty documented empirically in annotation research. Read and Carroll (2012) provide an instructive case, where two annotators independently coded [6]:

[6] Like him, Vermeer, or so he chose to believe, was an artist neglected and wronged by critics and who had died an almost unknown.

One annotator tagged *neglected* and *wronged* as Affect: Satisfaction, reading the passage as expressing Vermeer’s dissatisfaction with his treatment while another annotator tagged them as Judgement: Propriety, reading it as moral reproach directed at the critics, both readings semantically available. Read and Carroll (2012) acknowledge that the ambiguity reflects an insufficiently specified annotation guideline rather than an error by either annotator.

In the present annotation scheme, the Vermeer case is classified as Judgement: Propriety. AEs *neglected* and *wronged* do not foreground Vermeer’s felt emotional reaction; there is no explicit mood or surge, and he is not described as feeling distressed, angry, or sorrowful. Instead, the passage foregrounds the ethical status of the critics’ conduct: being wronged implies a violation of norms of fair treatment that holds independently of any individual’s emotional response. The following example from the present corpus provides a clear contrast:

[7] Some critics were frustrated that university officials opted to take the chance at all, given the state’s ongoing struggle to contain the virus. (20-08-US-NT)

(AS: *some critics*)
(AE: Affect: Dis/satisfaction: *frustrated*)
(AT: *university officials opting to take the chance*)

AE *frustrated* directly names an emotional state experienced by AS *critics*. The evaluation is explicitly anchored in their subjective experience; the frustration is the evaluative content, not an inference about anyone’s moral conduct. There is no assessment against the dimensions of social esteem or social sanction; the clause simply reports that a group of individuals experienced a negative emotional feeling.

The evaluative weight of *wronged* in [6] above lies in the ethical but not emotive domain. The contrast between Examples [6] and [7] yields the annotation guideline: when the clause foregrounds what or how someone feels as its primary evaluative content, annotate AE as Affect; when AE foregrounds how someone behaves, assessed against the dimensions of social esteem or social sanction, annotate AE as Judgement, even if the expression carries emotional resonance.

4. Decision Framework for Judgement versus Appreciation Classification

The Judgement-Appreciation boundary has attracted sustained attention in the Appraisal literature, but existing proposals have not converged on a resolution. Bednarek (2009) observes that grammatical patterns alone cannot reliably distinguish Judgement from Appreciation,

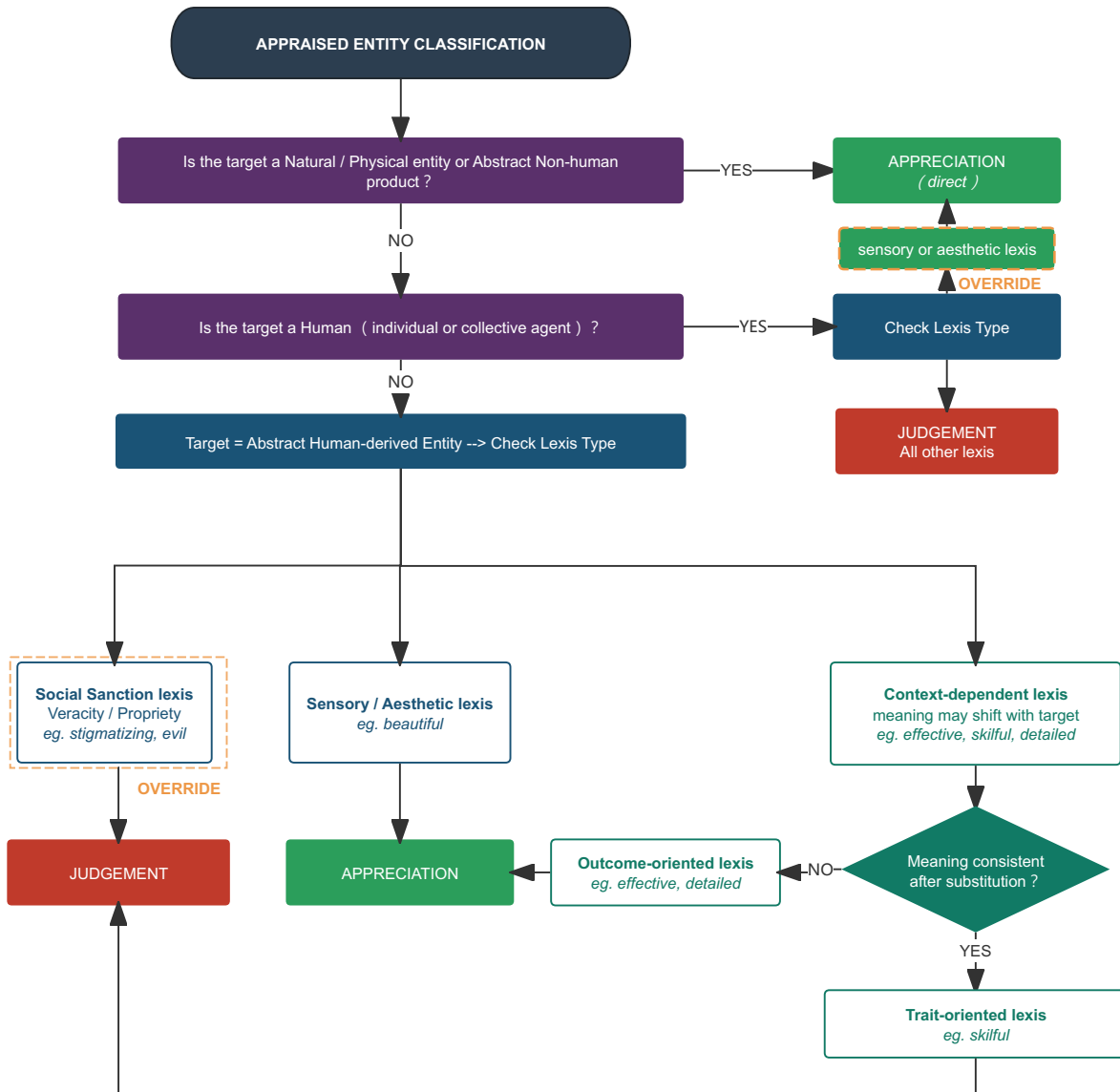


Figure 3: The complete decision framework for Judgement-Appreciation classification

proposing two key classification criteria, lexical type and target entity, that jointly determine category assignment without establishing clear precedence between them. Thompson (2014) prioritises target characteristics, suggesting that all evaluations of non-human targets should be classified as Appreciation, even when employing judging lexis, including nominalisations as non-human entities. However, Fuoli (2018) points out practical limitations: expressions like *industry-leading* and *disciplined* resist classification under traditional Appreciation subcategories while clearly evaluating human-related qualities. Starfield et al. (2015) address this issue by relocating Judgement subcategories into Appreciation: Valuation, accommodating judging lexis applied to non-human targets by expanding the Appreciation taxonomy. While this approach has resolved some

difficult classification issues, it has unnecessarily over-broadened the Valuation category beyond its original conceptual boundaries. Taboada & Carretero (2012) propose a different strategy, prioritising lexical properties through an ethics-aesthetics distinction: ethics-based evaluations belong to Judgement while aesthetics-based evaluations belong to Appreciation, regardless of target type, with abstract nouns classified as Judgement when representing nominalisations of human actions. In the present study, the framework described below adopts target type as the primary determinant of classification while refining the target taxonomy into three operationally distinct categories. Lexis type is retained as a secondary mechanism but its role is restricted to two well-defined override conditions and a substitution test, rather than operating as an independent co-equal

criterion. The decision framework for Judgement-Appreciation classification is visually represented in Figure 3.

4.1 Foundational principles

The framework rests on two interacting dimensions: the nature of the AT and the semantic core of lexis in AE. These two dimensions do not carry equal weight. AT type is the primary determinant of classification while AE lexis operates as a secondary mechanism that intervenes selectively through override conditions or through a substitution test. The central theoretical question underlying this priority order is the extent to which a human agent is implicated in the evaluation: Judgement evaluates human behaviour, character, and social conduct while Appreciation evaluates the properties of objects, phenomena, and artefacts.

AE lexis enters the framework as an override mechanism under exactly two conditions. First, AEs whose semantic core is inherently oriented towards social sanction dimensions, specifically Veracity or Propriety, trigger Judgement classification regardless of whether AT is human or nonhuman: the evaluative meaning presupposes human volitional agency and cannot be coherently detached from it. Second, AEs whose semantic core is inherently sensory or aesthetic in nature trigger Appreciation classification regardless of AT types since their evaluative meaning pertains to perceptual qualities rather than normative assessments of conduct or character. Outside these two override conditions, AT type governs classification throughout.

AEs whose semantic core does not pre-determine classification are designated in the present framework as context-dependent, in the sense that their evaluative meaning shifts according to the nature of AT. A substitution test is required when AEs are directed at abstract human-derived entities: The appraisal is transferred from the human-derived entity to its recoverable human agent, and the evaluative meaning of the resulting expression is examined for consistency. If the meaning remains essentially the same, that is, the evaluation retains the same qualitative dimension after substitution, AE is annotated as Judgement and retrospectively identified as trait-oriented AE, i.e. elements whose semantic core describes inherently human qualities that transfer coherently from non-human entities to human agents. If the meaning changes, that is, the evaluation shifts to a qualitatively different dimension after substitution, AE is classified as Appreciation and identified as outcome-oriented, i.e. elements whose semantic core describes functional effectiveness or compositional properties that cannot transfer to a human agent without altering the evaluative meaning. The substitution test therefore operationalises a distinction that is latent within the context-dependent AE category itself.

4.2 Classification Procedures

Category 1: Natural and physical entities, and abstract non-human products → Appreciation

Where AT refers to a natural or physical object, or an abstract entity not traceable to deliberate human action or decision, AE is classified as Appreciation without any further test. Human agency is either entirely absent or insufficiently implicated to anchor an ethical or social evaluation.

- [8] Two hotels in the exclusive Swiss Alpine resort of St Moritz were quarantined and ski schools closed to try to curb an outbreak of the highly infectious new coronavirus variant. (21-01-UK-TG)

{AS: implicit author}
{AE: Appreciation: *infectious*}
{AT: *coronavirus variant*}

The Delta variant is a natural entity. No human agency is implicated in its properties. AE *infectious* describes a biological characteristic of the variant itself, and the element is annotated as Appreciation.

- [9] Five days later, the outbreak was out of control, with nearly 400 virus cases among a campus student population that is usually around 6,000. (20-09-US-NT)

{AS: implicit author}
{AE: Appreciation: *out of control*}
{AT: *outbreak*}

AT *outbreak* is an abstract non-human entity, a developing epidemiological situation not directly traceable to a single deliberate human decision. AE *out of control* characterises the state of the outbreak itself and receives straightforward Appreciation classification.

Category 2: Human and collective human targets → Judgement

Where AT is human or functional as human, such as organisations, institutions, and government bodies, AE receives Judgement classification as the default outcome. Human agency is directly and fully implicated, satisfying the foundational condition for Judgement.

- [10] The island has struggled to secure vaccine supplies but the United States delivered 2.5 million Moderna doses last weekend. (21-06-HK-SM)

{AS: implicit author}
{AE: Judgement: Capacity: *struggled*}
{AT: *island*}

In [10], AT functions as a collective human agent responsible for its own public health governance. AE is therefore classified as Judgement: Capacity, reflecting a negative assessment of the collective

agent's ability to fulfil its responsibilities. AS is implicit, assumed to be the writer-speaker.

The sole exception to the human-target default arises when AE carries an inherently sensory or aesthetic semantic core. In such cases, the override condition takes precedence over the AT-primary rule and AE receives Appreciation classification regardless of the human AT. For instance, an evaluation such as *she is beautiful* describes a perceptual quality of the human AT's appearance whose semantic core cannot be coherently transferred to an evaluation of behaviour, competence, or social conduct; the Judgement reading is overridden and Appreciation assigned to AE. The human AT does not trigger the default Judgement outcome.

Category 3: Abstract human-derived entities

The operationally most complex category involves entities produced through or traceable to deliberate human action or decision such as reports, figures, plans, accusations, measures, and similar products of human agency. Classification in this category is determined by AE type, proceeding through three distinct steps described below.

Step 1: Social sanction AE override → Judgement

AE receives Judgement analysis if its semantic core is inherently oriented toward Veracity or Propriety. This override applies because the evaluative meaning presupposes a human agent whose conduct is being assessed against normative standards, regardless of whether AT is human. Consider the following example:

[11] The WHO chief stressed that having a name matters to prevent the use of other names that can be inaccurate or stigmatizing. (20-02-CN-CD)

⟨AS: WHO chief⟩
⟨AE: Judgement: Veracity: *inaccurate*⟩
⟨AT: other names⟩

⟨AS: WHO chief⟩
⟨AE: Judgement: Propriety: *stigmatizing*⟩
⟨AT: other names⟩

In [11], AT is *other names*, a human-derived entity. However, AE *inaccurate* evaluates the truthfulness of the naming practice: to call a name inaccurate is to assert that whoever produced it failed to represent reality truthfully, an evaluation that presupposes a human agent capable of truth or falsehood. AE *stigmatizing* evaluates the ethical conduct behind the naming: to stigmatise is to cause harm through social labelling, an evaluation that presupposes volitional human action subject to ethical norms. Neither evaluation can be coherently detached from human agency. The social sanction override applies to both items, which are classified as Judgement without proceeding to the substitution test.

Step 2: Context-dependent AE → Substitution test

The substitution test is applied in cases of AEs not covered by the social sanction or sensory-aesthetic overrides, where AT is transferred from the human-derived entity to its recoverable human agent, and the evaluative meaning of the resulting expression is examined for consistency.

The logic of the test can be illustrated through a pair of contrasting cases before turning to corpus examples. Consider

[12] It was a skilful innings.

The example is taken from Martin and White (2005, p. 59), where AE *skilful* is annotated as Appreciation. However, when the substitution test developed in the present framework is applied, substituting *a skilful innings* with *he is skilful* yields an evaluation with essentially the same meaning: both formulations assess the technical mastery. The consistency of evaluative meaning indicates that *skilful* is trait-oriented and therefore classified differently as Judgement in the present framework. Now consider AE *important*, which is noted in Bednarek (2009) as Appreciation when modifying a thing, issue, or proposition but as Judgement when modifying a person for the social standing and influence. The shift in evaluative meaning identifies *important* as context-dependent whose classification depends on the nature of the target.

These two contrasting cases illustrate the distinction that the substitution test is designed to operationalise. If the evaluative meaning remains essentially the same after substitution, the item is trait-oriented and receives Judgement classification, thus

Meaning consistent → Judgement (trait-oriented AE)

Consider [13], an example from the corpus of news discourse.

[13] One, by the name Kou, said: "With such a perfunctory state of emergency, holding the Olympics will be impossible." (21-01-SG-ST)

⟨AS: Kou⟩
⟨AE: Judgement: Tenacity: *perfunctory*⟩
⟨AT: state of emergency⟩

AT *state of emergency* is a human-derived entity traceable to a governmental decision. Substituting *a perfunctory state of emergency* with *the government was perfunctory* yields an evaluation of essentially the same meaning: both formulations assess diligence and commitment. AE *perfunctory* describes a disposition, a lack of seriousness or effort that transfers coherently from the product to its human agent. The evaluative meaning is consistent, and so AE is classified as Judgement.

If the evaluative meaning changes, the item is outcome-oriented and receives Appreciation classification, thus

Meaning inconsistent → Appreciation (outcome-oriented AE)

Now consider [14] as a final example in this study.

[14] A detailed study in California found that the variant easily spread from an unvaccinated teacher to children and, in a few cases, their families. (21-08-US-NT)

(AS: implicit author)

(AE: Appreciation: Composition: *detailed*)

(AT: *study*)

AT *study* is traceable to human action. However, substituting a *detailed study* with *the researchers are detailed* changes the evaluative meaning: the evaluation shifts from assessing the study's structural properties to a characterological assessment of the researchers' disposition. The change in meaning revealed through the substitution test classifies *detailed* as Appreciation.

5. Conclusion

This study has developed a methodological framework for the attitudinal annotation of evaluative language. It is grounded in Appraisal Theory and aims to address three persistent problems in the classification of evaluative language.

The first contribution is the three-component annotation model incorporating AS, AT, and AE for every evaluative instance. The model restores the relational structure of evaluation that is theoretically central to Appraisal Theory but has been routinely omitted in annotation practice. The selection principle developed alongside this model provides a principled procedure for resolving cases in which a single expression participates in more than one evaluative reading, replacing the ad hoc treatment of what Thompson (2014) termed the 'Russian doll' syndrome with a transparent decision mechanism grounded in the realisation and ontological status of AT.

The second contribution is the two-pathway model of the Affect boundary. By considering the relationship between Affect and the other two subsystems as two directional processes, where objectification tends toward Appreciation and socialisation toward Judgement, the model transforms the schematic notion of 'institutionalised Affect' discussed in Martin (2000) into two operational criteria. The first criterion distinguishes Affect from Appreciation by locating expressions along a gradient of experiencer salience, from individual felt emotion through social emotion to fully objectified property. The second criterion distinguishes Affect from Judgement by considering whether the expression foregrounds a

personal emotive state or an evaluation of conduct against the dimensions of social esteem or social sanction.

The third contribution is the decision framework for the Judgement-Appreciation boundary. The framework establishes the target type as the primary determinant of classification and defines two lexis-based override conditions (Social Sanction lexis and sensory or aesthetic lexis) that operate independently of the target type. For context-dependent lexis directed at abstract human-derived entities, the substitution test provides a falsifiable diagnostic: consistency of evaluative meaning after substitution indicates trait-oriented lexis and Judgement classification, while a shift in meaning indicates outcome-oriented lexis and Appreciation classification.

Several limitations should be acknowledged. The methodology has been developed and illustrated using English-language media discourse, and its applicability to other languages, registers, and genres remains to be tested. The selection principle addresses the most common configurations of multiple evaluative readings but does not claim to exhaust all possible cases; further corpus work may reveal configurations that require refinement of the priority ordering and related procedures. The substitution test, while designed to be replicable, involves an element of analytical judgement in determining whether the evaluative meaning has changed, and further empirical validation across corpora and annotation settings would be a valuable next step.

Despite these limitations, the methodology that we outline in this study offers a coherent and operationally explicit framework for Attitude annotation. By integrating the three-component model, the two-pathway criteria, and the target-lexis decision framework into a single analytical workflow, it aims to provide annotators with a sequence of principled decisions rather than a set of isolated category definitions. The objective has been to make the reasoning behind classification decisions transparent, replicable, and open to empirical testing so that disagreements in attitudinal analysis can be traced to specific points in the decision process rather than attributed to the inherent fuzziness of the categories themselves.

6. Acknowledgement

Research described in this article was partially supported by grants received from City University of Hong Kong (Project No 9360115), China National Social Science Fund (Project No 24&ZD28), and Beijing Social Sciences Foundation (Project Nos 18JDY005 and 19Y001).

7. Bibliographical References

Bednarek, M. (2009). Language patterns and

- ATTITUDE. *Functions of Language*, 16(2):165–192.
- Fuoli, M. (2018). A stepwise method for annotating APPRAISAL. *Functions of Language*, 25(2):229–258.
- Martin, J.R. (2000). Beyond exchange: APPRAISAL systems in English. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse*. Oxford University Press, pp. 142–175.
- Martin, J.R. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- Read, J. & Carroll, J. (2012). Annotating expressions of Appraisal in English. *Language Resources and Evaluation*, 46(3):421–447.
- Starfield, S., Paltridge, B., McMurtrie, R., Holbrook, A., Bourke, S., Fairbairn, H., Kiley, M., & Lovat, T. (2015). Understanding the language of evaluation in examiners' reports on doctoral theses. *Linguistics and Education*, 31:130–144.
- Taboada, M. & Carretero, M. (2012). Contrastive analyses of evaluation in text: Key issues in the design of an annotation system for attitude applicable to consumer reviews in English and Spanish. *Linguistics and the Human Sciences*, 6(1–3):275–295.
- Thompson, G. (2014). AFFECT and emotion, target-value mismatches, and Russian dolls: Refining the APPRAISAL model. In G. Thompson & L. Alba-Juez (Eds.), *Evaluation in context*. John Benjamins, pp. 47–66.
- White, P.R.R. (n.d.). An introductory tour through appraisal theory. Retrieved from <http://www.grammatics.com/appraisal/>
- White, P.R.R. (2004). Subjectivity, evaluation and point of view in media discourse. In C. Coffin, A. Hewings, & K. O'Halloran (Eds.), *Applying English grammar: Corpus and functional approaches*. Hodder Arnold, pp. 229–246.
- White, P.R.R. (2012). Exploring the axiological workings of 'reporter voice' news stories: Attribution and attitudinal positioning. *Discourse, Context & Media*, 1(2–3):57–67.
- Zeng, J., Wang, H., Bunt, H., Cao, X., Cardey, S., Dong, M., Hao, T., Jia, Y., Lee, K., Liao, S., Pustejovsky, J., Rey, F.C., Romary, L., Zong, J., & Fang, A.C. (2025). Enhanced evaluative language annotation through refined theoretical framework and workflow. In H. Bunt (Ed.), *Proceedings of the 21st Joint ACL–ISO Workshop on Interoperable Semantic Annotation (ISA-21)*, pages 76–84. Association for Computational Linguistics.