Exploring Moral Conflicts in Speech: Multidisciplinary Analysis of Affect and Stress

Minha Lee, Jaebok Kim[†], Khiet Truong[†], Yvonne de Kort*, Femke Beute* and Wijnand IJsselsteijn**
*Human-Technology Interaction
Eindhoven University of Technology
Eindhoven, The Netherlands
Email: {m.lee, w.a.ijsselsteijn, y.a.w.d.kort, f.beute}@tue.nl
[†]Human Media Interaction
University of Twente
Enschede, The Netherlands
Email: {j.kim, k.p.truong}@utwente.nl

Abstract—Our moral conscience as the "inner light" that guides us shines brighter during moments of ethical conflicts, when we notice a tension between our many oughts and/or wants. We present the first analyses on speech related stress and affect in accounts of moral conflicts. For our exploratory study, we started with interviews on moral and immoral events at work with entrepreneurs. Qualitative analysis revealed that interviewees do share personal moral conflicts without researchers probing for them. Quantitative analysis showed quiet and even toned voice features when discussing moral conflicts, and speech was laced with emotively positive and negative words, though more negative words were used. Moreover, we find promising results on our automatic classification experiment using speech features. How and what moral conflicts people deliberate on in real-life may be pertinent to future research in affective computing, as well as applications for decision-making support, ethical competences coaching, therapy, and healthy moral selfhood.

1. Introduction

An awareness of one's moral conflicts is a sign of a working conscience—a vulnerable moment before future regrets or worse, moral disengagement [1]. Moral conflicts are intrinsic to human psychology, yet their core features as observable phenomena are barely, if at all, investigated. As the first study, we aim to distill elements of moral conflicts via speech data. Technology that takes heed of *how* people dwell on real-life moral dilemmas can potentially aid in sudden lapse of judgement and pave the way for wiser deliberations. Since morally relevant events occur in everyday life [2][1], developing ethical expertise at a workplace [3] and supporting overall moral well-being through potential applications should rely on accounts of how people process moral conflicts. This could further help shape emotionally intelligent non-human entities, like robots or virtual agents,

to be sensitive to our moral dimensions if they are introduced as our coaches or companions.

Our current work is in line with our continuing investigation on moral stress. We frame our research on (1) a data set of authentic accounts of moral conflicts, and (2) analyses of accompanying affective speech features. Since this is the first time moral conflicts are examined as a unique phenomenon linguistically, i.e. - what is said, and paralinguistically, i.e. - how it is said, we build on a mixed-method approach of qualitative and quantitative analyses. In this study, we neither induced stress nor probed for moral conflicts artificially. So far, life-or-death moral conflicts are central to moral decision-making research [4], as well as task-based stress induction in stress detection [5]. We therefore looked into moral conflicts that originate from real-life experiences.

Our qualitative analysis revealed signs of psychological stress and moral conflicts regarding moral and immoral events at work. Thus for this paper, semi-structured interviews are our basis for exploring the question, *how are naturally occurring spoken accounts of moral conflicts unique in terms of linguistic and paralinguistic features?* We expected moral conflicts to show emotional responses both linguistically and paralinguistically, yet we mainly find linguistic signs of emotion, surprisingly both negative and positive affective words, though more negatively valenced words were present.

In section 2, we cover relevant literature on morality, affect, and speech processing, followed by section 3 on our methodological synthesis and data collection, and section 4 is on the analysis, results, and discussion. We summarize that linguistic and paralinguistic features of speech can potentially identify speech segments that cue moral conflicts. We conclude with suggestions on future research and final remarks in section 5.

2. Related work

Picard noted in 1995 that "recent neurological studies indicate that the role of emotion in human cognition is

---

1. Participants (N = 1,252) marked that there were morally relevant events 28.9% of the time in the past hour when they were randomly asked 5 times daily for 3 days on their phones [2].

essential. Emotions are not a luxury. Instead emotions play a critical role in rational decision-making [...]" [6]. Similarly, emotions are implicated in moral decision-making since early 2000s [4], [7], [8]. Moral conflicts may relate to both emotions and rational thinking. Albeit psychologically stressful, moral conflicts can be meaningful moments to deliberate on one's moral selfhood.

2.1. Moral decision-making and emotions

Various threads of research on human morality study moral judgments, reasons, and emotions, and the interplay between the three [4], [7], [8]. Moral reasoning comes before emotions in reaching moral judgments if a rational cognitive process is emphasized [9], [10]. However, emotions may precede moral judgments, followed by post-hoc reasoning to defend one's initial emotions [7], [11]. Moral reasoning and emotions can simultaneously shape our moral judgments as a dual process [12].

Research employing moral vignettes demonstrated that emotions are critical in moral decision-making. People with damaged ventromedial prefrontal cortex (VMPC), an area that mediates social emotions, were more likely to make a utilitarian[2] decision in what is known as the footbridge dilemma [8]. This is when participants are asked if they would push someone from a footbridge to block a runaway trolley to save five bystanders. In fMRI studies on participants without neural injuries, areas associated with emotion processing were activated when the footbridge dilemma was presented [4].

Manipulating emotions immediately before decision-making may affect decisions. For example, watching a humorous video before being presented with the footbridge dilemma reportedly increased positive affect and the likelihood of making a utilitarian choice [13]. Also, the longer participants spent on pondering over the footbridge dilemma, the more likely they were to choose a utilitarian act [4], [13]. Cognitive processing is stronger than emotional processing in both non-moral scenarios and moral scenarios requiring indirect involvement, such as flipping a switch to kill someone to save five people, while emotions are more at play in situations requiring direct involvement, such as pushing one person to death to save others [4]. Both emotional and cognitive aspects modulate moral decision-making to different degrees depending on the level of of personal involvement and contextual framing [12], [13].

2.2. Moral conflicts and the self

We define a moral conflict as a friction between incompatible moral obligations or principles [14]. Put differently, a moral dilemma occurs when there are alternatives that do not override one another on ethically relevant matters [15]. Alternative actions may all vie against each another as the most significant, right choice. Reflecting on real moral

conflicts can better prepare people to deal with future moral issues, both personal and societal dilemmas [15].

How one socializes impacts one's moral self identity, a central aspect of self-identity [16]. The moral self-image may be more essential to one's self-image than other markers, i.e. - autobiographical narrative or memory [17]. An awareness of a moral conflict may threaten one's established moral image [1], [14]. One may reinterpret what is ethical to preserve a consistent moral self-image, leading to moral disengagement by justifying one's actions as morally irrelevant, morally permissible, or even morally praiseworthy [1].

2.3. Verbal and non-verbal cues of moral conflicts

Moral conflicts involve indecisive positions, as well as discrete and non-discrete emotions, introducing worthwhile challenges for affective computing. In deception detection, inconclusive opinions had lowered accuracy than conclusive opinions in feature extraction; participants' responses on abortion were conclusive as in they were either for or against it, whereas when asked to truthfully talk about their best friends, participants' language was peppered with both negative and positive declarations, reducing precision due to inconclusive emotion categories in vocabulary choice [18]. Only linguistic analysis had equal accuracy for both deceptive and truthful conditions on the best friend topic [18]. The difficulty in deception detection due to mixed emotions resonates with our view of moral conflicts that embraces inconclusive judgments and emotions.

Self vs. other centeredness in processing social situations may be related to pronoun usage. When exposed to neutral pictures as probes, people diagnosed with post-traumatic stress disorder (PTSD) used more first-person singular pronouns, e.g. - I, and third-person singular pronouns, e.g. - she, but less third-person plural pronouns, e.g. - they. The future recurrence of PTSD related symptoms was strongly correlated to the usage of third-person singular pronouns [19]. Similarly, pronouns show a focus on victimhood, for people tended to rely more on first-person pronouns when describing being teased, and used more third-person pronouns when talking about teasing someone else [20]. Furthermore, an increase in emotional pain heightened the use of first person singular pronoun [20]. We thus looked into pronouns usage in moral conflicts, in case of a possible a trend.

Psychological states of stress do not always equate to detectable signs of stress in speech. Hence, comprehensive characteristics of speech, including prosodic, spectral, linguistic and nonverbal features, have been investigated. Especially, prosodic and spectral features, e.g. F0, energy, Zero-Cross-Rate (ZCR) and spectrogram, are known to be strongly associated with affective states of speakers [21], [22]. Also, acoustic characteristics reflecting contrastive stress constantly varied with affective states over semantically and syntactically identical utterances [23]. In a similar sense, when a speaker is under stress, vocal tract movement is affected, resulting in a quantifiable perturbation in positions of articulators. In order to study speech under cognitive

---

2. A typical utilitarian act attempts to do the greatest amount of good for the greatest number of people.

and emotional stress on the vocal tract spectrum, formant locations and bandwidth have been examined [24]. Speech under stress is known to cause the widest shift in F1 formant locations. Particularly, bandwidth of F1 and F2 significantly varies in terms of mean while that of F3 and F4 does not. The variance of formant location and bandwidth also shows shifts [24].

Other non-verbal aspects of speech, e.g. speaking rate, duration of speech and pause, were also examined in relation to stress. Acute stress reduced fluency in speech when participants performed the Trier Social Stress Test (TSST) [5]. Interestingly, participants spoke with more pauses and with increased vocabulary productivity during TSST compared to non-stressful speech [5]. Quantitative analysis and subjective reports of stress "in the wild" at a call center showed that person-specific notion of stress is a useful direction because individuals experience and display stress in different ways [25]. Furthermore, moral stress at call centers foretold lowered work satisfaction, job fatigue, and higher turnover intent [26]. Combating work-related moral stress may be a valuable aim.

Linguistically, stress and negative affect influence people's word choices. We used Linguistic Inquiry and Word Count (LIWC) [27], which allows for word usage analysis at scale and across a broad range of signals touching on linguistic markers of cognition and affect. LIWC is often used to analyze the usage of affective words such as happiness and sadness, and it is extensively validated. In therapeutic settings, using negatively valenced words during reflection may signal signs of future improvements, helping people to cognitively process and emotionally release negativity for psychological health betterment [28].

3. Data

We employed a methodological fusion of qualitative research and quantitative speech analysis. This is based on the notion that both qualitative and quantitative research methods are non-neutral, and combining both is valuable for showing unexpected trends [29], especially when it comes to a novel subject matter, such as moral conflicts. Thus, the first qualitative thematic analysis of interviews on moral events led us to confront that moral conflicts are conversationally brought up.

Our coded segments were manually selected by annotators, and we concentrated on speech segments on moral conflicts that were extracted. Most interviews in qualitative research are recorded and transcribed, but audio files can also be used to examine paralinguistic features, serving a dual-purpose as we have utilized. This enhanced qualitative analysis and provided rich data for speech analysis, resulting in greater multidisciplinary insight. From the onset, the point of the interviews were not to look for moral conflicts, for the study behind interview questions is not about moral conflicts, but about occurrences of moral and immoral behavior in everyday life [2]. Hence, the theme emerged during interviews in a natural manner and participants were not primed to think about moral conflicts because the researchers themselves were unaware of moral conflicts to begin with. Only after data collection moral conflicts as a theme was detected.

3.1. Data collection

Our data consists of interviews on moral and immoral acts at work, as well as morally irrelevant topics. Our interviewees were entrepreneurs or freelancers. This specific population was of interest for qualitative research to observe a possibly unique type of moral stress. Successfully navigating both personal and collective interests is crucial to entrepreneurship [30], yet this duality of being selfish while being considerate of the greater good may bring moral stress for many entrepreneurs. Moral stress, not moral conflict, was the motivating factor to begin the research process.

We had 22 interviewees, recruited via snowball sampling at accelerators, co-working spaces, and maker-spaces in the Netherlands. Two sessions were excluded from analysis. One case failed to meet the criterion of not mentioning the word "stress" before the interview; the interviewer purposefully did not mention stress before or during all interviews to not prime participants to think about stress. The other case was excluded due to a technical failure when an interviewee gesticulated with a mic. We thus have 20 sessions, totaling approximately 15 hours. Of the included interviewees, half of them were non-Dutch, nine were women. 17 participants provided their age and the average was 32.5 years *(SD = 7.4)*. Interviews were in English and recorded with a portable omni-directional mic with a sampling rate of 16khz. All interviews were conducted by the first author and they took place at interviewees' work locations, except one participant who requested to do the interview at home.

Before all semi-structured interviews, participants signed the informed consent form that neither mentioned moral conflicts nor stress, but morality in general. The starting topics were on their work environments and routines. Afterwards, there was a series of eight questions on morally good and immoral acts that were witnessed, heard about, committed (participant as the doer) or experienced (participant as the recipient), categories lifted from [2]. There were subquestions on experienced emotions and further elaborations. Lastly, participants were asked if they still think about discussed events and about the proudest moments of their careers. After the interview, participants were debriefed that the recorded material will undergo speech and qualitative analysis. None of the participants chose to withdraw their consent.

3.2. Annotation

Interviews were manually transcribed based on audio material, and then transcripts and audio were entered to ELAN[3], where annotations on moral conflicts and question-answer pairs were coded for speech processing analysis.

---

3. ELAN - http://www.mpi.nl/corpus/html/elan/

Categories	High awareness	Low awareness
Excerpts	*I come from a family where they do discriminate. They don't know they do, but they really do. And it's really hard not to be that way because I'm raised that way. I always really want to try hard not to do that. So if someone does that I feel really awkward. [...]I don't feel the same way as this person does or says [...], I completely disagree, but it's hard to say or do something about it because I feel awkward about it, because my family does it as well. I never know what to do about it.*	*[...] here I also see tricking, and you know it's not really like people are very generous. Some are! I wouldn't say nobody is, but it's like I found out that some people say "I don't know this", "I'm competent in this" and it's actually not true and everything is a mess like this. And it's about the art of saying those things...*
Explanation	**The participant (female, age 24) has an ongoing moral conflict on workplace discrimination. She shows high level of awareness because she admits that discriminatory acts cause her to feel conflicted without a definite course of action.**	**The interviewee (female, age 39) corrects herself to share that there are generous people around her, even though she commented that people are not generous at her workspace in the previous sentence. She perceives others' possibly dishonest sayings as disingenuous generosity, but she does not directly comment on this and how it influences her, which classifies her position as low awareness.**

TABLE 1. *Examples of moral conflicts. Both quotes were annotated as moral conflicts by coders.*

For qualitative analysis we used RDQA[4]. We focused on interviewees' responses as either moral conflicts or non-moral conflicts. Based on interviews, a moral conflict is defined as (1) either a past event or an ongoing issue of a moral nature that one still thinks about or has recently reflected on, and/or (2) with no definite conclusion on how to address those concerns, as in no consistent emotions and/or judgments when discussing ongoing or past issue(s). This means participants expressed mixed emotions or mixed judgments regarding an event, for example, feeling both disappointed and relieved.

Interviewees may display varying levels of awareness in criterion two, with high awareness involving a direct comment on being conflicted with at least two diverging views, and low awareness consisting of contradictory remarks on the topic, without the speaker commenting on how there are opposing ideas on the same subject. A contradiction can be made during a single moment in an interview, or at two (or more) different points during the interview on the same topic. We instructed our annotators to code interviewees' speech with the above description of a moral conflict and examples are located in Table1. Whether or not a discussed topic is firsthand morally relevant is in accordance with [2].

Currently, moral conflicts have been identified by both coders. The first annotator coded all sessions. Our second annotator coded five interviews based on the above criteria of moral conflicts, and then we checked for reliability. Moral conflicts are difficult to code, for annotators are more likely to code any segment as non-moral conflicts since they occur more frequently; moral-conflict segments are sparse. Since we did not seek out moral conflicts, an equal balance between moral conflict and non-moral conflict segments could not be reached by the nature of how qualitative interviews were conducted. Thus, we compute Gwet's AC1 [31] that is robust against high-prevalence, and we obtained .749 ($p < .01$), which is moderately high. We concluded to use only segments only from the primary coder due to the high reliability and the sparseness of moral

conflict segments. The main point here is to demonstrate that moral conflicts in speech should be looked at more closely without first assuming that directly asking about moral conflicts or artificially inducing moral conflicts is the best possible starting point in data collection. When dealing with uncharted territories such as moral conflict recognition, performing data collection with classifier training in mind should be a secondary step, after recognizing that a phenomenon researchers want to classify exists based on studies that prioritizes participants' viewpoints as a starting point, not researchers' prior assumptions that already frame how participants should behave or think.

Only the primary coder looked into speakers' level of moral conflict awareness because we do not have enough data on moral conflict awareness to always treat it as a discrete element of low and high, when it may be considered as a spectrum. Due to the small amount of data on moral conflict awareness, we present it here as a critical factor to look into when more thoroughly studying moral conflict in the future. Of the 132 identified moral conflict segments by the primary coder, 77 were noted as high awareness and 55 as low awareness. Speech segments per speaker were topically dependent, meaning there were not 132 unique moral conflicts in total, but 132 points in speech about moral conflicts. For speech analysis, 87 moral conflict segments remained because segments that are shorter than 5 sec. and longer than 1 min. were removed in case of misleading analysis due to temporal inconsistencies.

4. Analysis and Results

4.1. Qualitative Analysis

According to our participants, moral conflicts have fluid boundaries in terms of self involvement. Most conflicts revolved around immoral acts involving team members or co-working community members, but attribution of blame is difficult to pinpoint, for many participants see themselves as entangled in most work events that are memorable. Other-centered moral conflicts were described with emotive

4. RQDA - http://rqda.r-forge.r-project.org/

phrases like "disgusted" to describe the low awareness situation in Table 1, to milder expressions like "miscommunication" to not label a conflict as morally relevant, though this participant (male, 31) "would not trust [a former colleague] anymore" after the long-term "communication problem".

Critical evaluations of one's own behaviors were reported less often, but possibly more emotionally complex. To list key aspects, self-centered moral conflicts involved feelings of *self-hypocrisy* when one's actions are not aligned with one's own moral values and of one's business, *polarization* when high-stakes of running a startup is both psychologically gratifying and exhausting, *insecurity* when one sees one's arguments or position as weak, and *disappointment*, *confusion*, *anger*, and/or *defensiveness* when one's actions are perceived as unethical by others or by oneself, even though one believes that that may not be an accurate interpretation.

Contemplations on moral misdeeds can create empathy and increase understanding. For instance, the participant who demonstrated high awareness in Table 1 shared that "[discrimination] is not always meant in a bad way [...]. People make mistakes and of course I do too. But you also understand other people more when you realize you did that yourself". Another participant (male, 27) also noted that "sometimes it is just human, human(s) make mistakes. [...] I tried not to be a hypocrite on that (being treated unfairly in a competition). I wont say I will never do that, but if I would be in that kind of position (a judge) I would try to look back [...]. It might be just a small decision for us, for me to maneuver some candidates, but it could be a life changing decision for them. So, [...I would] at least rethink a bit". Addressing moral conflicts over immoral acts committed by oneself and others are chances to re-evaluate one's values and to be less judgmental in human fallibility. Surprisingly, an awareness of moral conflicts may be an opportunity to sympathize with perceived immoral acts, which in parallel becomes a reflective trigger for moral self betterment.

4.2. Quantitative Analysis

Our qualitative analysis has shown that moral conflicts do come up without an interviewer probing for them, and that conflicts are more complex than simple distinctions between immoral and morally good events. In this section, we substantiate how these complex distinctions impact speech in a quantitative way and share results of classification experiments on moral-conflict and non-moral conflict segments in speech.

4.2.1. Speech features. We looked into a comprehensive set of features based on related works [24] since we do not have pre-knowledge of features of moral conflicts. We explored three different feature sets: non-verbal, linguistic, and prosodic (and spectral) features, as summarized in Table 2. We applied statistical functionals on the features depending on the categories. For the prosodic and spectral features, we utilized various functionals such as mean, max, min, standard-deviation (SD) of values, slopes, and more. We

Category	Features
non-verbal features	speech and pause
linguistic features	positive and negative words, personal (possessive) pronouns
prosodic & spectral features	Fundamental frequency (F0), energy, Zero-Cross-Rate (ZCR), formant (F1, F2, F3, F4), Harmonic-Noise-Ratio (HNR), jitter, shimmer

TABLE 2. FEATURE SETS

Category	Speech		Pause	
	Count	Duration (s)	Count	Duration (s)
moral-conflict	685	$4.13 \pm .38$	851	$.560 \pm .52$
non-moral-conflict	4703	$4.46 \pm .39$	5126	$.542 \pm .49$
total	5388	$4.45 \pm .39$	5977	$.544 \pm .49$

TABLE 3. *Descriptive statistics of non-verbal features*

used the eGEMAPS [32] feature set to cover this wide range of variants. For linguistic features, we counted the number of used affective words and personal possessive pronouns.

We automatically extracted speech segments from the labelled moral-conflict and non-moral conflict segments via Voice Activity Detection (VAD). VAD was configured to extract speech based on "Inter-Pausal Units (IPUs)" (1.0 sec of the minimum length for speech and pause) of speech [33]. We bridged two successive speech segments only if they were separated by a short silence (< 1.0 sec). Based on IPUs, we obtained features for subsequent analyses. From each IPU, we acquired prosodic and spectral features by using the Opensmile toolkit [34] and a configuration of eGEMAPS [32]. Linguistic features were identified using Empath, a sentiment analysis tool [35]. Empath covers a broader range of categories than LIWC [27]; it is capable of building lexicon mined from texts on the web and adding categories, using a combination of deep learning and crowd-sourcing. In our case, we focused on positive and negative emotional categories.

4.2.2. Results: verbal and non-verbal analysis of speech. We started with overall patterns of non-verbal aspects of speech. We collected all speech and pause segments for each interviewee and calculated the number and mean duration of the segments. Table 3 summarizes the results. To assess differences of the features between the categories, we transformed feature values to z-scores for each interviewee and conducted the Kruskal Wallis test ($df = 1, p < .01$) [36]. Overall we had a greater number of non-moral-conflict than moral-conflict speech segments, and the duration of moral-conflict speech segments was significantly shorter ($p < .01$) than that of non-moral conflict speech segments. In addition, the duration of moral-conflict pause segments was longer than non-moral conflict pause segments, but the difference was not statistically significant ($p > .01$).

We analyzed prosodic and spectral features and then we extracted features that were transformed into z-scores for each interview to consider individual differences. Afterwards, we categorized the features into three: F0, loudness, and formants. Table 4 shows the results that have a signifi-

Features	F0		
	Freq-SD	Rising-SD	Falling-SD
moral-conflict	-.103	-.038	-.076
non-moral-conflict	+.013	+.005	+.010
	Loudness		
	M	Rising-M	Falling-M
moral-conflict	-.099	-.006	-.066
non-moral-conflict	+.013	+.007	+.008
	Formants		
	F1-Freq-SD	F1-Band-SD	F2-Freq-SD
moral-conflict	-.043	-.090	-.036
non-moral-conflict	+.006	+.011	+.005

TABLE 4. *Prosodic and spectral features; Freq (Frequency), Band (Bandwidth), Rising (Rising slope), Falling (Falling slope), M (Mean), SD (Standard-deviation)*

cant ($p < .001$) difference between moral-conflict and non-moral-conflict segments. First, the frequency of F0 was less variant in speech on moral-conflict compared to that of non-moral-conflict. Moreover, rising and falling slope of F0 were less variant in speech of moral-conflict. Hence, changes of F0 seemed to be flatter in speech in moral-conflict segments. Next, loudness of speech in moral-conflict was smaller than that in non-moral conflict sections. Also, rising and falling slope of loudness was smaller in moral-conflict than those in non-moral-conflict. Therefore, interviewees tended to speak softly and keep their volume at the same level. Participants showed smaller variances of frequency and bandwidth of formants 1 and 2 in moral-conflict condition, in contrast to speech under cognitive stress [24]. In short, interviewees spoke less loudly and kept flat tones. This does not fully align with previous research on speech under stress, which has detected vocal qualities to be louder with higher tones and more variations, e.g. - [24].

moral-conflict				non-moral-conflict			
pos (.016)		neg (.024)		pos (.001)		neg (.001)	
word	prop	word	prop	word	prop	word	prop
better	22.0	bad	20.0	better	13.2	mean	21.1
feeling	18.3	mean	17.5	happy	11.5	bad	14.2
trust	12.2	hard	15.8	proud	11.5	hard	14.2
family	11.0	angry	5.8	feeling	11.1	thinking	8.7
proud	9.8	care	5.8	family	10.3	either	6.8
care	8.5	wanted	5.0	great	9.8	care	6.2
keep	7.3	thinking	4.2	care	8.5	reason	5.3
happy	6.1	fault	4.2	trust	8.1	wanted	4.6
wish	6.1	surprised	3.3	honest	7.7	angry	4.3
hope	6.1	alone	3.3	reason	7.3	stop	2.5

TABLE 5. *Frequencies of affective words and proportion (%) of top-10 words; pos (positive), neg (negative), prop (proportion)*

Additionally, we investigated affective words. We tokenized all sentences and took out stop words by using a lexicon and Part-Of-Speech (POS) dictionaries. For example, we did not include articles (a, an, the). Then, we counted positive and negative words through Empath and normalized the words by the total number of spoken words per speaker. Table 5 shows the mean values (M) of the normalized frequencies for each category and proportion of top 10 words. There were more affective words in moral-conflict segments than non-moral-conflict segments. In moral-conflict segments, there were more negative words

	first singular	third singular	third plural
moral-conflict	.185	.140	.252
non-moral-conflict	.160	.311	.310

TABLE 6. *Frequencies of personal pronouns*

($M = .024$) than positive words ($M = .016$). This relates to [18], for both positive and negative affective words can be used to truthfully describe a real-life, personal opinions, i.e. - on one's best friend. In non-moral-conflict segments, there was no difference between the number of positive ($M = .001$) and negative words ($M = .001$). When we ordered the top 10 words by proportion to the total number of affect-laden words, there were only few different words depending on the categories. We believe that open questions resulted in these similar patterns of using affective words.

Lastly, we covered frequencies of personal (possessive) pronouns as displayed in Table 6. Though the difference was not large, we found more first-person singular pronouns in moral conflicts, a possible sign of emotional involvement as per [20]. In non-moral conflict segments, more third-person singular and plural pronouns were observed at an equal rate; we saw less other-centeredness in moral conflicts, suggesting that moral conflicts were more self-focused. This diverges from how those with PTSD use more first and third-person singular pronouns, and less third-person plural pronouns [19]. But again, experimental paradigms make comparisons difficult; [19] asked participants to freely talk about a picture, a neutral probe, and we relied on qualitative interviews which are more like bi-directional conversations, not uni-directional narratives. It will be beneficial to revisit linguistic patterns, including pronoun usage, when moral conflicts and non-moral conflicts in speech are more balanced in speech data.

4.2.3. Automatic classification of moral conflict and non-moral conflict speech. We assessed the feasibility of automatically detecting moral conflicts by conducting feature selection using Information Gain Ratio (IGR) [37] and Random Forest classification, an ensemble learning method for classification by building a multitude of decision trees [38]. We sampled 685 segments from non-moral conflict speech for classification experiment because of the severely imbalanced distribution of moral ($N = 685$) and non-moral conflict speech segments ($N = 4703$). Table 7 presents top 10 features. The top 10 features were prosodic & spectral features. We found that only prosodic & spectral features were observed up to top 30 features. Other features such as affective linguistic cues did not show a high IGR. We proceeded with a classification experiment using the top 30 features. Because of the sparseness and imbalanced distribution of samples over our interviewees, we did not conduct Leave-One-Speaker-Out-Cross-Validation. Rather, we shuffled all samples and split them into 10 folds in a random manner and ran 10-fold cross-validation. The results are in Table 8. Overall F-score was significantly higher than chance at .647. Although this static modelling did not capture temporal dynamics of the speech (the mean duration: 4.13 ± 38),

IGR	feature
.0054	slopeUV0.500-M
.0043	hammarbergIndexUV-M
.0036	alphaRatioUV-M
.0036	mfcc4-SD
.0035	F0-Falling-SD
.0035	slopeV500.1500-SD
.0032	mfcc3-SD
.0032	mfcc3V-SD
.0031	Mean-Unvoiced-Segment-Length
.0030	HNRdBACF-SD

TABLE 7. *Top 10 features selected by IGR, Rising (Rising slope), Falling (Falling slope), M (Mean), SD (Standard-deviation)*

the selected feature set achieved potential performance. We see opportunities for further improvement using temporal models in the future.

	moral conflict	non-moral conflict
moral conflict	453 (66%)	232 (34%)
non-moral conflict	263 (38%)	422 (62%)

TABLE 8. *Confusion matrix. number of segments (portion)*

4.3. Discussion

We comprehensively looked into moral conflicts through qualitative and quantitative analyses of spoken accounts. Moral conflicts matter because ethical issues may fester due to unresolved overlapping priorities, and possibly lead to moral disengagement. One must first acknowledge a moral conflict in order to search for solutions; moral conflict detection can bring matters to the forefront of personal or organizational awareness. Despite the fact that we did not probe for moral conflicts directly, we see promising signs of moral conflicts as a distinguishable phenomenon via qualitative interviews, and quantitatively at 64.7% for uni-modal automatic classification. Our current work shows that people speak more calmly with less variance in speech when discussing morally conflicting matters and use affectively positive, and to a larger extent, negative words. We thus posit that linguistically, moral conflicts involve emotive elements, and paralinguistically we see pensive tendencies. This can be interpreted in more than one way. Perhaps it is a sign of proactive moral engagement, an attempt at avoidance of cognitive dissonance, or simply moments of level-headed contemplation, while using affect driven vocabularies; emotions are recalled without being relived vicariously.

While dwelling on moral conflicts without a definite conclusion may be stressful psychologically, it is difficult to conclude with this exploratory study that interviewees are experiencing stress according to established methods, for research on detection of stress in speech has relied mostly on cognitive stress tasks and simulated stressful situations [24], induced acute forms of stress [5], or recognizably stressful natural scenarios [25]. Moral decision-making research is equally difficult to fully relate to, for it thus far heavily leans on speculative vignettes such as the trolley dilemma

[4], [8]. Overall, an awareness of one's own moral conflicts is a sign of morally engaged reflections that may require paradigmatically different interpretations.

4.4. Future Works

Future research should employ more data collection on moral conflicts directly, and apply other analysis techniques in speech processing, as well as looking into multi-modal moral conflict classifications. With a data set consisting of a more balanced number of segments on moral conflicts and non-moral conflicts, we can train a classifier based on greater input from annotators as a logical step forward. Moral conflict detection can benefit a wide range of decision-making support, such as applications for clinical, military, or policy related issues. For personal well-being and digital therapy, conflict detection in narrative accounts can alert people to revisit taxing matters or seek professional help. In organizational settings, identifying moral conflicts can aid in ethical expertise development and help to avoid systematic numbness to ethical cues. Embedding digital assistants, like affective virtual agents or robots, with moral conflict detection abilities can help them react to human moral nuances and/or to help train people for empathic awareness. We foresee research on moral conflicts applied to wider domains, broadening our learnings beyond targeted stress or emotion recognition and working towards mediating technologies; framing moral saliency is important for ameliorating, disambiguating, and emphasizing moral conflicts that personally and societally matter to us. Research on moral conflicts can show affective computing as a powerful supporting actor in many fields, which in turn may redefine it as an independent domain.

5. Conclusion

Our work is methodologically and topically novel, for we synthesized qualitative and quantitative research to explore moral conflicts for the first time. The data was based on interviews on moral and immoral events at work. After the initial qualitative screening that detected moral conflicts as a phenomenon, we analyzed linguistic and and paralinguistic features of morally conflicting parts of speech compared to morally irrelevant and morally non-conflicting speech components. We found that both negatively and positively valenced words were more often used in speech on moral conflicts than in non-moral conflicts, though there were more negative affective words. Paralinguistically, interviewees spoke quietly with less variance in tone during morally conflicting talks. Yet whether or not participants are undergoing stress or negative affect whilst talking about morally conflicting events is inconclusive; we may need more comprehensive interpretations of stress, affect, and moral decision-making.

The rationale behind looking into moral conflicts is that they are integral to our lives, be it a conflict on being a whistle blower at the cost of losing one's job or breaking a self-promise to stand up for equality by choosing to uphold

personal safety. We believe that acknowledging moral conflicts is healthy for both individuals and our greater society. All humans err and how we process our moral and immoral acts shape conscientious decision-making. Our exploratory contribution on moral conflicts introduces many avenues for multidisciplinary research.

References

[1] A. Bandura, *Moral Disengagement: How People Do Harm and Live with Themselves*. New York: Macmillan, 2016.

[2] W. Hofmann, D. C. Wisneski, M. J. Brandt, and L. J. Skitka, "Morality in everyday life," *Science*, vol. 345, no. 6202, pp. 1340–1343, 2014.

[3] E. Dane and S. Sonenshein, "On the role of experience in ethical decision making at work: An ethical expertise perspective," *Organizational Psychology Review*, vol. 5, no. 1, pp. 74–96, 2015.

[4] J. D. Greene, R. B. Sommerville, L. E. Nystrom, J. M. Darley, and J. D. Cohen, "An fMRI investigation of emotional engagement in moral judgment," *Science*, vol. 293, no. 5537, pp. 2105–2108, 2001.

[5] T. W. Buchanan, J. S. Laures-Gore, and M. C. Duff, "Acute stress reduces speech fluency," *Biological psychology*, vol. 97, pp. 60–66, 2014.

[6] R. W. Picard, "Affective computing," Technical Report 321, MIT Media Lab, Perceptual Computing Group, 1995.

[7] J. Haidt, "The emotional dog and its rational tail: a social intuitionist approach to moral judgment." *Psychological review*, vol. 108, no. 4, p. 814, 2001.

[8] M. Koenigs, L. Young, R. Adolphs, D. Tranel, F. Cushman, M. Hauser, and A. Damasio, "Damage to the prefrontal cortex increases utilitarian moral judgements," *Nature*, vol. 446, no. 7138, pp. 908–911, 2007.

[9] I. Kant, *Kant: The metaphysics of morals*. Cambridge University Press, 1996.

[10] R. J. R. Blair, "Moral reasoning and the child with psychopathic tendencies," *Personality and individual differences*, vol. 22, no. 5, pp. 731–739, 1997.

[11] D. Hume, *A treatise of human nature*. Courier Corporation, 2003.

[12] J. Greene and J. Haidt, "How (and where) does moral judgment work?" *Trends in cognitive sciences*, vol. 6, no. 12, pp. 517–523, 2002.

[13] P. Valdesolo and D. DeSteno, "Manipulations of emotional context shape moral judgment," *Psychological science*, vol. 17, no. 6, pp. 476–477, 2006.

[14] B. Monin, D. A. Pizarro, and J. S. Beer, "Deciding versus reacting: Conceptions of moral judgment and the reason-affect debate," *Review of General Psychology*, vol. 11, no. 2, p. 99, 2007.

[15] W. Sinnott-Armstrong, *Moral dilemmas*. Wiley Online Library, 1988.

[16] K. Aquino and A. Reed II, "The self-importance of moral identity." *Journal of personality and social psychology*, vol. 83, no. 6, p. 1423, 2002.

[17] N. Strohminger and S. Nichols, "The essential moral self," *Cognition*, vol. 131, no. 1, pp. 159–171, 2014.

[18] M. Abouelenien, V. Pérez-Rosas, R. Mihalcea, and M. Burzo, "Deception detection using a multimodal approach," in *Proceedings of the 16th International Conference on Multimodal Interaction*. ACM, 2014, pp. 58–65.

[19] S. Papini, P. Yoon, M. Rubin, T. Lopez-Castro, and D. A. Hien, "Linguistic characteristics in a non-trauma-related narrative task are associated with ptsd diagnosis and symptom severity." *Psychological Trauma: Theory, Research, Practice, and Policy*, vol. 7, no. 3, p. 295, 2015.

[20] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: Liwc and computerized text analysis methods," *Journal of language and social psychology*, vol. 29, no. 1, pp. 24–54, 2010.

[21] M. El Ayadi, M. S. Kamel, and F. Karray, "Survey on speech emotion recognition: Features, classification schemes, and databases," *Pattern Recognition*, vol. 44, no. 3, pp. 572–587, 2011.

[22] A. Vinciarelli, M. Pantic, and H. Bourlard, "Social signal processing: Survey of an emerging domain," *Image and Vision Computing*, vol. 27, no. 12, pp. 1743–1759, 2009.

[23] L. Cosmides, "Invariances in the acoustic expression of emotion during speech." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 9, no. 6, p. 864, 1983.

[24] J. Hansen and S. Patil, "Speech under stress: Analysis, modeling and recognition," *Speaker classification I*, pp. 108–137, 2007.

[25] J. Hernandez, R. Morris, and R. Picard, "Call center stress recognition with person-specific models," *Affective computing and intelligent interaction*, pp. 125–134, 2011.

[26] K. B. DeTienne, B. R. Agle, J. C. Phillips, and M.-C. Ingerson, "The impact of moral stress compared to other stressors on employee fatigue, job satisfaction, and turnover: An empirical investigation," *Journal of Business Ethics*, vol. 110, no. 3, pp. 377–391, 2012.

[27] J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic inquiry and word count: Liwc 2001," *Mahway: Lawrence Erlbaum Associates*, vol. 71, no. 2001, p. 2001, 2001.

[28] J. W. Pennebaker, "Writing about emotional experiences as a therapeutic process," *Psychological science*, vol. 8, no. 3, pp. 162–166, 1997.

[29] A. Bryman, "Integrating quantitative and qualitative research: how is it done?" *Qualitative research*, vol. 6, no. 1, pp. 97–113, 2006.

[30] A. H. Van de Ven, H. J. Sapienza, and J. Villanueva, "Entrepreneurial pursuits of self-and collective interests," *Strategic entrepreneurship journal*, vol. 1, no. 3-4, pp. 353–370, 2007.

[31] K. L. Gwet, *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. Advanced Analytics, LLC, 2014.

[32] F. Eyben, K. R. Scherer, B. W. Schuller, J. Sundberg, E. André, C. Busso, L. Y. Devillers, J. Epps, P. Laukka, S. S. Narayanan *et al.*, "The geneva minimalistic acoustic parameter set (gemaps) for voice research and affective computing," *IEEE Transactions on Affective Computing*, vol. 7, no. 2, pp. 190–202, 2016.

[33] M. Heldner and J. Edlund, "Pauses, gaps and overlaps in conversations," *Journal of Phonetics*, vol. 38, no. 4, pp. 555–568, 2010.

[34] F. Eyben, M. Wöllmer, and B. Schuller, "Opensmile: the munich versatile and fast open-source audio feature extractor," in *Proceedings of the international conference on Multimedia*. ACM, 2010, pp. 1459–1462.

[35] E. Fast, B. Chen, and M. S. Bernstein, "Empath: Understanding topic signals in large-scale text," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 4647–4657.

[36] N. Breslow, "A generalized kruskal-wallis test for comparing k samples subject to unequal patterns of censorship," *Biometrika*, vol. 57, no. 3, pp. 579–594, 1970.

[37] A. G. Karegowda, A. Manjunath, and M. Jayaram, "Comparative study of attribute selection using gain ratio and correlation based feature selection," *International Journal of Information Technology and Knowledge Management*, vol. 2, no. 2, pp. 271–277, 2010.

[38] T. K. Ho, "Random decision forests," in *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, vol. 1. IEEE, 1995, pp. 278–282.