

Perceptions related to engaging in non-driving activities in an automated vehicle while commuting: A text mining approach

Yilun Xing¹, Linda Ng Boyle², Rafaela Sadun³, John D. Lee⁴, Orit Shaer⁵, and Andrew Kun⁶

^{1,2}University of Washington, Seattle, WA

³Harvard University, Boston, MA

⁴University of Wisconsin-Madison, Madison, WI

⁵Wellesley College, Wellesley, MA

⁶University of New Hampshire, Durham, NH

ABSTRACT

Automated vehicles provide the human operator an opportunity to engage in non-driving tasks while traveling. In this study, we examined and compared drivers' perception of non-driving tasks in a future with highly automated vehicles with that of current vehicle systems, where the human operator needs to control braking and steering. The study used a survey that was distributed through an online paid marketplace platform called Lucid. The survey included open-ended questions soliciting participants' perceptions of non-driving tasks given a work commute scenario for each driving mode. Text mining and cluster analysis were used to analyze the responses from 752 participants in four open-ended survey questions. Results showed that drivers had higher positive sentiments towards future AVs than towards current vehicle systems. The most-reported non-driving activity was "listen" for current vehicle systems and "work" for AVs, though "work", "listen", and "relax" were the most reported overall. The change in drivers' perception from current systems to AV

systems was also captured. The findings show that most drivers (83.4%) would not change their non-driving activities, with 76.0% continuing to perform work or work-related activities. Approximately 8.7% of respondents reported that they would abandon their previous tasks to do work or work-related tasks in AV, while 3.7% reported the opposite—they would abandon work-related tasks to do other tasks. The results suggest that working while commuting will be an advantage of AVs. Hence, understanding how those in the car can work productively will be important as we move forward with automated vehicles.

Keywords: Automated Vehicles, Text mining, Commuting, Non-Driving Tasks

1 INTRODUCTION

Drivers tend to divide their attention and engage in multiple activities while driving (Alt et al., 2010). In a survey by McEvoy et al. (2006), drivers reported a range of “distracting activities” while driving; this included lack of concentration (71.8%), adjusting in-vehicle equipment (68.7%), viewing outside people, objects or events (57.8%), talking to passengers (39.8%), drinking (11.3%), eating (6.0%) or smoking (10.6%). The emergence of automated vehicle (AV) technology can support the human operator in performing the driving task so they can be more productive and safe on their non-driving activities during their travel.

Knowledge workers are individuals who gather, analyze, interpret, and synthesize information to advance understanding of specific subject areas and provide knowledge to organizations to make better decisions (Frick, 2010). Given their job responsibilities, knowledge workers place great importance on productivity and may have a stronger need to accomplish work-related tasks during their commute. They are often paid according to their productivity rather than the number of hours in repetitive work tasks (Drucker, 1999). Unlike manual workers which require production equipment for their job (Drucker, 1999), knowledge workers usually have greater flexibility in their working environment. These differences showcase the benefits that automated vehicles may have for knowledge workers, which makes them more likely to leverage automation for productivity outside an office or home environment.

While AVs may allow drivers to divide their attention between driving and non-driving tasks, there is still the potential for increased crash risk (Bálint et al., 2020) and badly designed interfaces can increase motion sickness (Young et al., 2007) and undermine the opportunity to relax. Further, the balance between home and work life can negatively impact a knowledge workers’ overall well-being. To effectively provide drivers the opportunity to perform non-driving tasks in AVs, we need to first understand people’s perception of conducting non-driving tasks while in the car. Past studies on perceptions of AVs are often limited to a user’s overall perceptions of AV rather than

1 their preference or ability to engage in other non-driving tasks (Hudson et al., 2019; Milakis et al.,
2 2017; Hulse et al., 2018; Penmetsa et al., 2019). In one study conducted by Cyganski et al. (2015)
3 on potential activities in AV, music and talking to passengers were reported to be the most often
4 conducted activities while driving. Their study showed that working while traveling played a minor
5 role and was rarely considered as an advantage of AVs by participants.

6 The goal of our study is to understand people's perception of and willingness to engage in
7 non-driving activities while in an automated vehicle when compared to current manual driving.
8 Our focus is on trips related to the daily commute. This is the most frequent trip among workers
9 (Malokin et al., 2019) and often fosters engagement in non-driving activities. Studies showed that
10 commuting on long trips can be productive and enjoyable (Humagain and Singleton, 2020). There
11 may also be different non-driving activities depending on whether the individual is commuting to
12 work or home. For example, the commute home may provide more time to think about non-work-
13 related or relaxation activities (Varghese and Jana, 2018). People's desire to be productive while
14 commuting and the possibility that AVs will free the drivers partly or fully is the motivation for this
15 study. This study addresses three research questions:

- 16 1. What changes are expected in the drivers' general attitude towards performing non-driving
17 tasks while commuting?
- 18 2. What changes are expected between non-driving tasks that are non-work and work-related?
- 19 3. How will drivers' perception towards performing non-driving tasks change when imaging
20 using AVs for commute?

21 **2 DATA SOURCE**

22 A time-use survey was distributed using the on-line platform Lucid, which partners with several
23 companies to recruit individuals to answer online surveys (A copy of the survey is available at

(Teodorovicz et al., 2021)). The purpose of the survey was to understand how knowledge workers use their time during a typical work day, with a focus on performing non-driving activities in AVs. We selected two open-ended questions, with each question including two parts. The open-ended questions were used to solicit responses related to the perception of commuting in their current vehicle and the possibilities with an AV. The question asked participants to imagine a situation where they would be commuting in fully automated vehicles (Reja et al., 2003). Participants were required to answer the following two questions for two difference scenarios (morning commute, evening commute); there was no word limit.

Q1. Currently, how and why do you currently engage in non-driving tasks when commuting in the morning/evening?

Q2. Now, picture yourself in a technologically advanced future. Imagine one of these advanced technologies is a safe self-driving car. How would you use it? What and why would you do while commuting?

There four responses were categorized as:

1. Current Morning: commute in the morning with current vehicle systems.
2. Current Evening: commute in the evening with current vehicle systems.
3. AV Morning: commute in the morning with future AVs.
4. AV Evening: commute in the evening with future AVs.

Lucid received \$13 per sample collected and the team did not have control over how much of this value is each participant compensated (Teodorovicz et al., 2022). Exposing detailed information such as autonomous vehicles may attract participants who have preference over the topic. Therefore, the purpose of the survey was always advertised as “to understand time use, and how that affects productivity and well being.” Potential participants were screened for three criteria

1) employed in a full-time job at the time of participation (+35 hours/week); 2) earning an annual

1 salary income of at least \$40,000 US dollars which corresponds to approximately the 6th percentile
2 of the income distribution of knowledge workers in the US; 3) working in a "knowledge worker"
3 occupation.

4 There were of 756 commuting knowledge workers that responded to the survey. Of these, four
5 was removed given incomplete information for a total of 752 used for the subsequent analysis.

6 Among the 752 participants, 346 are female and 406 are male. Participants' age ranges from 18
7 to 79 with an average of 38.02. Most participants (96.5%) resides in the US with 21 NA's and 5
8 from other counties such as France and India (the country of residence were mapped with lat/long
9 coordinate).

10 3 METHODS

11 We applied a text mining approach to examine the open-ended responses. There are two concepts
12 commonly used in text mining: *document* is a unit of analysis of textual data, which corresponds
13 to each response in a question (Ronen and James, 2007), and *terms* are the components that make
14 up the *document* (Ghazizadeh et al., 2014).

15 Figure 1 shows a flow chart of the analysis procedure. The data sources (Box 1) are cleaned for
16 analysis (Box 2). We begin with a sentiment analysis (Box 3) to examine drivers' general attitudes
17 towards commuting with AVs. Term frequency analysis (Box 4a) is then applied to identify the
18 most frequently occurring terms (e.g, read, work) that correspond to the non-driving tasks. Term
19 co-occurrence analysis (Box 4b) is then used to identify the distance between the target term and
20 other terms (e.g., work and email; work and scheduling).

21 The open-ended responses provide insights on drivers' non-driving tasks. We therefore exam-
22 ined the changes in perception toward non-driving tasks if AVs were available (e.g., listen to music
23 during morning commute in manual driving, to work during morning commute in AVs) (Box 5).
24 All responses were clustered into different term groups and each group was labeled with the most

1 frequent term in that group (Box 5a). Participants' responses were converted from text data to cat-
2 egorical data. Because each participant had four responses (or four commute conditions), we have
3 four independent variables for each participant, and these are used for the participant clustering
4 analysis (Box 5b). The clustering results show how each group's perception of non-driving task
5 engagement changed from the current condition to an envisioned future, when they will be able to
6 use AVs.

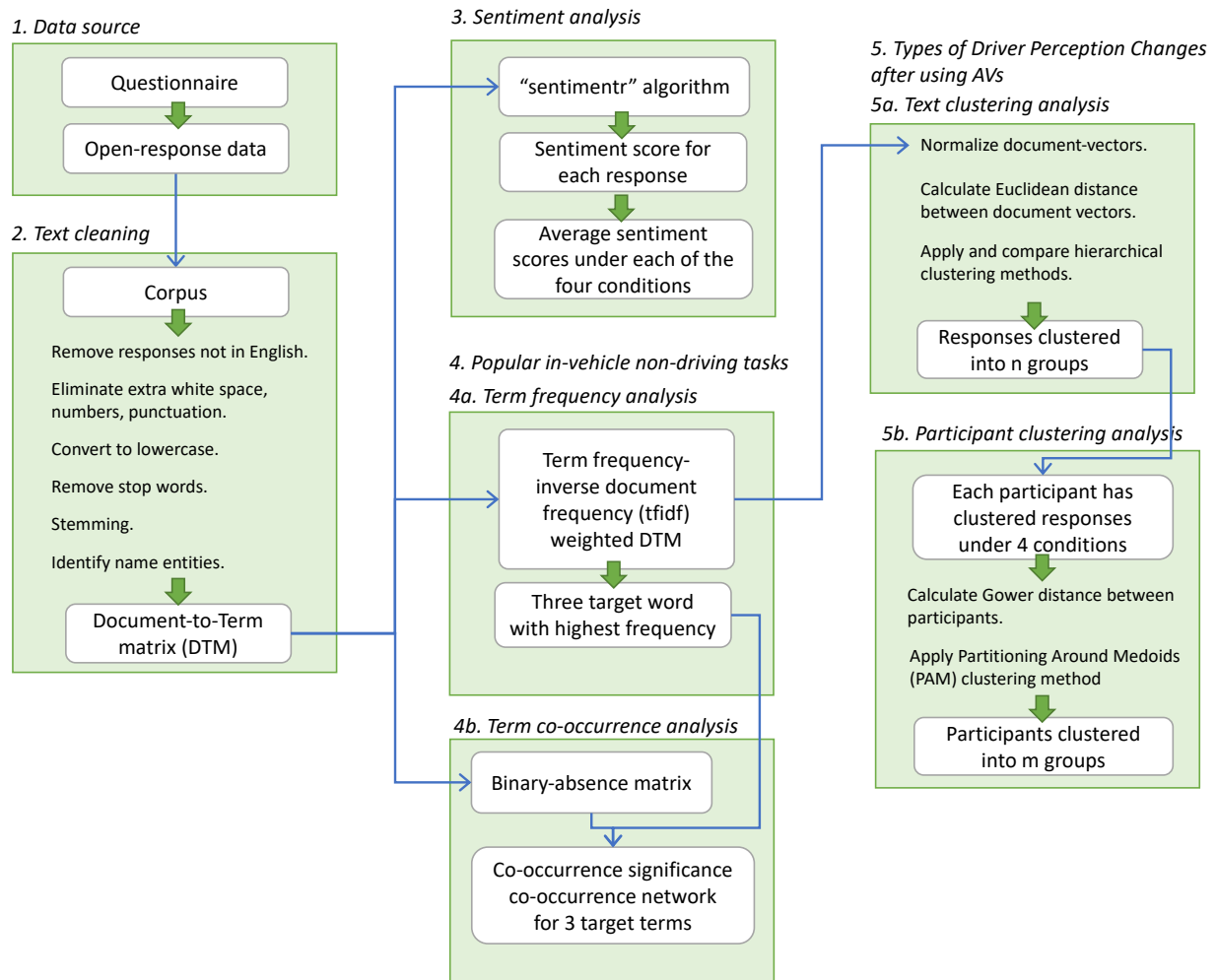


Figure 1. The analysis procedure of our study

3.1 Text Cleaning Process

The following seven steps are commonly used prior to text mining (Das et al., 2019; Ghazizadeh et al., 2014) and were customized for our data. The processes were carried out using the R programming language (version 4.1.0) using the packages `tm` (Feinerer, 2015) and `quanteda` (Benoit et al., 2018).

1. Convert text data to corpus data structure.

The corpus data structure is the data structure used for text analysis in R packages `tm` and `quanteda`. A corpus class object contains the original documents, document-level variables, document-level metadata, corpus-level metadata, and default settings for subsequent processing of the corpus .

2. Remove responses that were not in English.

Six responses were in another language and were removed. For example, “es muy util hacer todo” in Spanish.

3. Eliminate extra white spaces, numbers, punctuation and convert to lowercase.

Our analysis focused on term frequency, co-occurrence rather than semantic analysis, so the position or sequence of the words does not matter. Numbers were considered not meaningful without specific content. This step made sure that each word was examined separately.

4. Remove stop words.

Stop words are generally the most common words in a language and do not add value to the analysis. The `tm` package provides a list of 174 stop words for English (e.g, the, a, is) that can be accessed by function `stopwords(kind = "en")`. We also developed a list of custom stop words: “car,” “vehicle,” “drive,” “self-driving,” “non-driving,” “commute.”

1 5. Stemming.

2 Stemming is a method that converts words to their radicals, e.g., 'working' 'worked' 'works'
3 becomes "work".

4 6. Identify named entities.

5 Some words appear together frequently and represent a single concept. This includes phrases
6 such as "social media" "phone call". Words that appeared together ten or more times in the
7 dataset were combined into a single term with "_".

8 7. Transform to Document-to-Term matrix (DTM).

9 Corpus data were transformed to a DTM and then used for analysis. Each row of the DTM
10 refers to a document, and each column of the DTM refers to a term. Each entry of the DTM
11 represents the number of occurrence of a term in a document.

12 After the text cleaning process, the number of unique terms in our corpus for the Current Morn-
13 ing survey item was reduced from 854 terms to 632; Current Evening from 756 to 564; AV Morning
14 from 928 to 703; AV Evening from 847 to 641. The number of non-missing value for the Current
15 Morning and Current Evening survey responses reduced to 673; AV Morning to 696; AV Evening
16 to 610 (e.g., responses such as "no", "0" were removed through the processes). Therefore, for
17 the text clustering procedure, there were (696+696+673+610=) 2675 responses included. For the
18 participant clustering procedure, since a response for each of the four conditions is required, 574
19 participants' data were used as input of the clustering model. Among the 574 participants, 249 are
20 female and 325 are male. Participants' age ranges from 21 to 79 with an average of 38.0. Most
21 participants (97.0%) resides in the US with 14 NA's and 3 from other counties such as India (the
22 country of residence were mapped with lat/long coordinate).

3.2 Sentiment Analysis

The sentiment analysis was conducted using R package `sentimentr` v2.9.0 (Rinker, 2021). The package provides a good tool for addressing valance shifters which include *negator* (flips the sign of a polarized word, e.g., “I do *not* like it.” A polarized word refers to a positive or a negative word), *amplifier* (intensifier, e.g., “I *really* like it.”), *de-amplifier* (downtoner, e.g., “I *hardly* like it.”) and *adversative conjunctions* (overrules the previous clause containing a polarized word, e.g., “I like it *but* it’s not worth it.”). A review (Naldi, 2019) on R packages for sentiment analysis concluded that the `sentimentr` package performs best in addressing the valance shifters’ issues. All the four state-of-art packages compared (`syuzhet`, `Rsentiment`, `sentimentr` and `SentimentAnalysis`) adopt the bag-of-words approach, where the sentiment score is calculated based on the individual words in the text neglecting the role of syntax and grammar. Each package has their own lexicons for polarized words with or without negators. However, `sentimentr` is the only package that has a separate lexicon with 140 valance shifters and their weights. The other three packages compute the sentiment score by summing up the number or weight of all the polarized word in a sentence. Whether the negators are considered depends on whether the lexicon includes negators or whether the user search for the negators and reverse the sentiment score (i.e., multiplied by -1) with the number of negators being odd, e.g., “I do not like it” 1) will be mislabelled with a positive score without negators in the polarized word lexicon and the users’ additional work, 2) will reduce the weight of the negator from the sentiment score if negators are included in the polarized word lexicon and 3) will be reversed to have a negative score if users’ additional work is involved.

The `sentimentr` package addresses the valance shifter issue in a more comprehensive way. First, the words in each sentence are searched and compared to a lexicon of polarized words with each positive word tagged a +1 sentiment score and negative a -1 respectively. Then, a polarized context subset consisting of the polarized word, the four words before and two words after it, are pulled out from each sentence. These words around the polarized words are considered as the

1 valence shifters that will have an impact on the sentiment of the polarized word. The words in this
2 subset are tagged as neutral, negator, amplifier or de-amplifier. Each neutral hold no value, while
3 each valance shifter is assigned to a weight varied by the valence shifter type. Amplifiers become
4 de-amplifiers if the subset contains an odd number of negators. A detailed description on how the
5 weights are assigned can be seen in the package help file. The accumulated weight are then added
6 to the sentiment score of the polarized word. Last, these weighted polarized context subset scores
7 are summed and divided by the squared root of the word count yielding an unbounded sentiment
8 score for each sentence. The average sentiment score of each questionnaire response is to get the
9 mean of all the sentences within it. A positive sentiment score indicates an overall positive attitude
10 and the greater the value the more positive attitude the participant holds.

11 3.3 Term Frequency Analysis

12 Term frequency-inverse document frequency (*tfidf*) was first introduced by Jones (1972) and is
13 widely used for natural language processing (NLP). It uses the database size and the distribution of
14 terms to determine the weights. We first use the frequency of each term as its weight. Hence, terms
15 that appear more frequently are assumed to be more important and descriptive for the document.
16 Let $D = d_1, \dots, d_n$ be a set of documents and $T = t_1, \dots, t_m$. Let $tf(d_i, t_j)$ denote the frequency
17 of term $t_j \in T$ in document $d_i \in D$, then we have

$$18 \quad tf(d_i, t_j) = \frac{N(d_i, t_j)}{N(d_i)}$$

19 where $N(d_i)$ represents the number of terms in document i , and $N(d_i, t_j)$ represents the number of
20 term j in document i .

21 However, this is not usually the case in practice. The most frequent terms are not necessarily the
22 most informative ones; terms that appear frequently in a small number of documents but rarely in the
23 other documents tend to be more relevant and specific for that particular group of documents, and
24 therefore more useful for finding similar documents. In order to capture these terms and reflect their

importance, we transform the basic term frequencies $tf(d_i, t_j)$ into the $tfidf$ weighting scheme. $tfidf$ weighs the frequency of a term t_j in a document d_i with a factor that discounts its importance with its appearances in the entire document collection, which is defined as

$$tfidf(d_i, t_j) = tf(d_i, t_j) \times \ln\left(\frac{n}{df(t_j)}\right)$$

where n represents the number of documents in the corpus, and $df(t_j)$ (document frequency) represents the number of documents in which term t_j appears.

In this study, $tfidf$ were used for all the analysis except for the sentiment analysis that uses the exact number of positive/negative terms to calculate the proportion of each type terms.

3.4 Term Co-occurrence Analysis

Co-occurrence analysis seeks to explore the joint occurrence of terms in a context window, which can be documents, paragraphs, sentences, or neighboring terms as opposed to free term combinations (Kolesnikova, 2016). Each document (or open response) in our study was relatively short (only one to three sentences). The responses were also of a loose structure unlike those observed in organized blogs. Further, the terms' relationship with other terms in the document matters. Hence, the documents were not divided into sentences, which is typical process for long blogs, but rather used directly as the context window for the term co-occurrence.

The first step is to transform the DTM into a binary-absence matrix in which each row is a document and each column is a term. Each entry of the binary-absence matrix represents the presence (1) or absence (0) of a term in a document. A co-occurrence matrix (sometimes called a Term-Term-Matrix (TTM)) is calculated by multiplying the transposition of the binary-absence matrix and itself. Each entry represents the co-occurrence times of the corresponding pair of terms.

The significance of co-occurrence was calculated by the Dice statistic (Dice, 1945). It is a widely used association measure for detecting co-occurrence that is simple and shown to have outstanding performance among measures (Wiedemann and Wiedemann, 2016; Kolesnikova, 2016).

1 The formula used for calculating the Dice statistic of all responses A containing one term a and all
2 responses B containing one term b is:

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|} = \frac{2n_{ab}}{n_a + n_b}$$

4 where n_{ab} represents the number of joint occurrence of term a and term b , n_a represents the number
5 of occurrence of term a , and n_b represents the number of occurrence of term b .

6 The value of this index ranges from 0, which indicates a failure to detect association for the
7 pair of terms in our corpus, to 1 for a perfect association. Given the large number of pairings, a
8 co-occurrence network was applied to help visualize the joint occurrence relationship of terms (see
9 Figure 4 - 6). The network includes nodes and edges. Nodes represent the terms and the association
10 between terms. The network begins with a target term at degree zero. Each subsequent degree is
11 computed from an existing term in the network and linked with edges. Nodes with less than two
12 edges are removed.

13 3.5 Text Clustering

14 The term frequency method (*tfidf*) is shown to be as good as the semantic based methods (e.g.,
15 Latent Semantic Indexing) for text clustering (Schütze and Silverstein, 1997). *tfidf* was chosen
16 because it is more intuitive to tag each text cluster with their most frequent term first (e.g, work)
17 and then group participants based on their responses from the text cluster (goal of this study). The
18 semantic method, which shares similar ideas with Principle Component Analysis (PCA), represents
19 terms as a linear combination of other terms (Thomas et al., 1998). The findings are not as intuitive
20 as term frequency methods, making it more difficult to tag the clustered answers to a term.

21 Each row of the *tfidf* weighted DTM matrix is a vector representation of a document. Distance
22 between the two vectors was then measured to capture the similarity of the two documents and
23 clustered into groups based on this similarity. Euclidean distance was used as it is widely used
24 in clustering problems, including clustering text (Jing et al., 2006; Lee et al., 2012). To reduce

the computation cost and better compare the similarity of documents (reduce the impact of the document length), the vectors representing the documents were first normalized within the corpus using Euclidean norm:

$$\widehat{DTM}_i = \frac{DTM_i}{|DTM_i|}$$

where $\widehat{DTM}_i$ represents the normalized *tfidf* weighted vector for document i , DTM_i represents the *tfidf* weighted vector for document i , and $|DTM_i|$ represents the norm of the *tfidf* weighted vector for document i .

The Euclidean distance between each pair of documents was then calculated as:

$$Dist_E(d_{i_1}, d_{i_2}) = |\widehat{DTM}_{i_1} - \widehat{DTM}_{i_2}| = \sqrt{\sum_{j=1}^m [\widehat{DTM}_{i_1j} - \widehat{DTM}_{i_2j}]^2}$$

where $\widehat{DTM}_{i_1}$ and $\widehat{DTM}_{i_2}$ represents the *tfidf* weighted vectors for document $i_1 \in \{1 \dots m\}$ and $i_2 \in \{1 \dots m\}$ correspondingly. Function `hclust()` from R package `stats` was used for the text clustering. Four cluster methods "average", "centroid", "ward.D" and "ward.D2" were tested, and "ward.D2" (Murtagh and Legendre, 2014) performed best in generating meaningful hierarchical clusters.

3.6 Participant Clustering

Participants were clustered based on their response types generated from the text clustering analysis for the four open-ended questions. That is, four nominal variables (categorical, not ordered) were used for clustering the participants. Similarity between participants was measured using "Gower distance", which computes the average of all feature-specific distances. For categorical features, the distance equals zero if two participants have the same value; otherwise, the feature's value will be one. The function `daisy()` from R package `cluster` was used for calculating the Gower distance.

Partitioning Around Medoids (PAM) was used for clustering the participants' characteristics (e.g., work in morning with AV) based on the Gower distance Kaufman and Rousseeuw (2008).

This method is similar to k-means. In k-means, the centroids are identified using the mean of all data points assigned to that centroid's cluster. Medoids are the points that have minimum dissimilarity with other points within the clusters.

The procedures of the PAM method includes: 1) randomly selecting k data points as initial medoids; 2) Assign each data point to the closest medoids, based on distance; 3) Improve the quality of clustering by exchanging selected objects with unselected objects; 4) Repeat steps 2 and 3 until the positions of the medoids no longer change and the sum of distances of individual units from medoids is as small as possible. Silhouette method was applied to get the optimal number of clusters by examining the width of the clusters.

$$Dist.out(d_i) = \min_{d_j \notin C} (Dist(d_i, C))$$

$$s(d_i) = \frac{Dist.out(d_i) - Dist.in(d_i)}{\max(Dist.out(d_i), Dist.in(d_i))}$$

where, $Dist(d_i, C)$ represents the average distance of document i to all documents of cluster C to which it does not belong. $Dist.out(d_i)$ represents the average distance of document i with other documents in the same cluster. If document i is the only document in its cluster, the value will be zero. $dist.out(d_i)$ represents the minimum average distance of document i to all observations from another cluster.

A good clustering outcome would have high similarity within a cluster and high dissimilarity between clusters. Thus, optimal cluster numbers would be achieved when the silhouette coefficient reaches its maximum value.

4 RESULTS AND ANALYSIS

4.1 Sentiment Analysis

The sentiment score distributions under the four driving conditions are shown in Figure 2. There were more responses with positive sentiment scores than negative scores in all conditions, indicat-

1 ing an overall positive attitude towards doing non-driving tasks during commuting. Participants had
 2 a more positive sentiment toward doing non-driving tasks with AVs (mean=0.1376, SD=0.2899)
 3 than in the current driving condition (mean=0.1326, SD=0.3000); and a more positive sentiment
 4 toward doing non-driving tasks during morning commutes (mean=0.1448, SD=0.2989) than in the
 5 evening (mean=0.1254, SD=0.2907); No significant difference were detected at a 0.05 level via
 6 t-test.

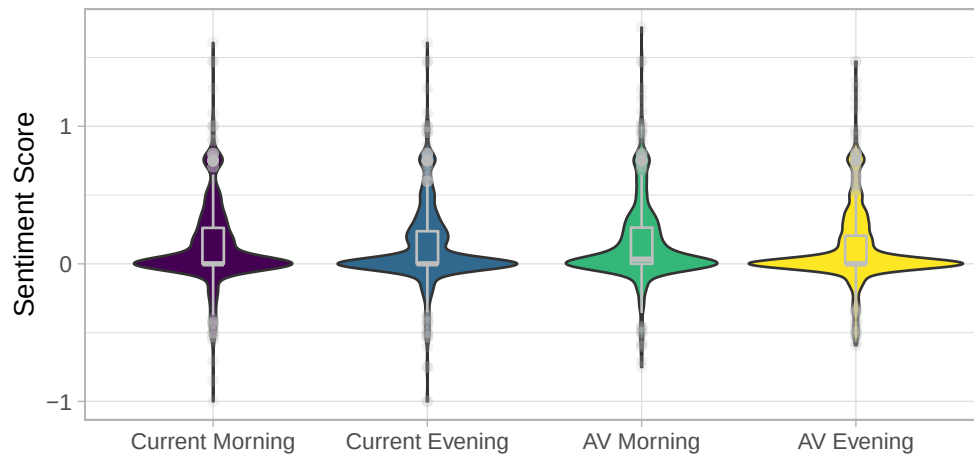


Figure 2. Violin plot for the sentiment scores for the four driving conditions

7 When comparing the sentiment among the four conditions, results showed a more positive
 8 sentiment for AV morning (mean=0.1536, SD=0.2962) then for current driving in the morning
 9 (mean=0.1361, SD=0.3015), than for current driving in the evening (mean=0.1292, SD=0.2986)
 10 than for AV in the evening (mean=0.1216, SD=0.2828). No significant difference was detected
 11 through pairwise t-test at the 0.05 level. Though the two AV conditions did not both have more
 12 positive scores than the two current driving conditions, there were less responses with more negative
 13 sentiment scores, i.e., few responses with a sentiment score below -0.5. Examples of sentiment
 14 score output are shown in Figure 3 with green highlights the positive and red highlights the negative.

+476

I would get much more time to do other activities like checking emails, receiving more phone calls, listening to music etc.

.000

None

+245

I would program it to get me efficiently to work and I would casually be aware of the trip but probably surf the web or facebook

-.177

I would not trust a self-driving car.

+208

Finish simple jobs that I couldn't do before due to lack of time

Figure 3. Examples of responses highlighted with green and red to show its polarity.

4.2 Popular In-Vehicle Non-Driving Tasks

4.2.1 Term Frequency Analysis

The top 20 terms in the *tfidf* weighted frequency for the four conditions (Current Morning/Evening; AV Morning/Evening) are shown in Table 1. The types of non-driving tasks conducted while commuting was different for morning and evening. People reported a preference toward work-related tasks in the morning and a preference toward relaxing or listening to radio/music in the evening.

This preference for non-driving tasks also differed between current driving and using an AV, which can be summarized as follows:

1. "listen" is the most frequent secondary task with current vehicle systems while "work" is most frequent for AV.
2. "phone" is more often used with current vehicle systems.
3. drivers preferred to "read" in an AV

4. drivers preferred to "sleep" in an AV, which is not detected as a prevalent task for the current vehicle.

Table 1. Top 20 terms with the highest *tfidf*. "Listen," "work," "relax" are highlighted to show their ranking change among conditions.

Rank	Current Morning		Current Evening		AV Morning		AV Evening	
	Word	TF-IDF	Word	TF-IDF	Word	TF-IDF	Word	TF-IDF
1	listen	86.3	listen	76.7	work	95.8	relax	69.1
2	work	64.8	music	56.1	time	71.5	work	64.1
3	time	62.9	radio	55.4	read	69.3	time	60.0
4	radio	62.0	work	53.5	get	65.0	read	55.7
5	music	56.1	time	49.6	relax	54.5	listen	50.8
6	good	54.1	phone	45.1	sleep	48.8	get	41.8
7	phone	52.1	relax	44.7	day	47.4	home	40.9
8	get	47.9	good	41.9	good	45.6	good	39.9
9	none	40.9	none	40.9	listen	43.8	none	33.7
10	relax	35.8	home	37.8	probably	41.9	personal	33.0
11	like	35.2	like	30.3	will	39.3	music	32.6
12	engage	34.8	talk	30.3	like	36.9	sleep	31.4
13	task	33.0	get	29.6	thing	36.3	new	31.4
14	need	33.0	can	28.4	emails	35.8	day	31.2
15	day	31.9	go	27.9	take	35.8	will	31.2
16	podcasts	31.4	traffic	25.9	prepare	34.1	take	30.7
17	go	29.1	activity	25.4	make	34.1	probably	30.3
18	traffic	28.4	think	25.4	new	33.7	like	30.3
19	just	26.7	engage	24.2	none	32.6	check	29.1
20	can	25.4	make	24.2	email	31.4	even	28.8

Based on these findings, it seems AVs will change the types of tasks drivers engage in while commuting. The non-driving tasks appear to change from auditory to visual (the extreme case is sleeping). Hence, attention will be directed away from driving when people commute using an AV. Individuals may also transition toward other in-vehicle devices beyond phones. The detailed ranking of the top 20 terms in *tfidf* from Table 1 show that checking or editing emails and preparing for the day rank higher in AVs and will most likely require larger screens and keyboards.

Three top-ranking terms were chosen for further analysis: "listen", "work", and "relax". These three targets are highlighted in Table 1, and each of the terms for the different conditions are linked

1 together to visualize their change in frequency. The frequency of “work” and “relax” will exceed
2 “listen” when AV is used for commuting.

3 4.2.2 Term Co-occurrence Analysis

4 Based on the term frequency analysis, “work”, “relax” and “listen” were selected as target terms
5 for further examination. Term co-occurrence analysis was used to investigate how people work,
6 relax and listen under the current and AV situations. In co-occurrence analysis, the target term is
7 placed in the center of a network and the relationships to other nodes are examined (See Table 2).

Table 2. Significance measured in Dice statistics from the target terms (work, relax, and listen) to the top 10 corresponding secondary terms/nodes.

Rank	from Work				from Relax				from Listen			
	Current		AV		Current		AV		Current		AV	
	to	sig	to	sig	to	sig	to	sig	to	sig	to	sig
1	home	0.22	get	0.24	help	0.23	read	0.19	radio	0.58	music	0.53
2	get	0.17	day	0.18	can	0.20	day	0.15	music	0.47	radio	0.42
3	go	0.11	time	0.17	listen	0.19	time	0.12	podcasts	0.27	podcasts	0.27
4	take	0.11	prepare	0.14	stress	0.17	spend	0.11	relax	0.15	podcast	0.18
5	always	0.10	relax	0.12	away	0.17	enjoy	0.11	time	0.15	still	0.15
6	day	0.10	read	0.12	prefer	0.16	work	0.10	keep	0.12	probably	0.13
7	stress	0.09	personal	0.11	music	0.15	listen	0.10	like	0.12	check_emails	0.12
8	time	0.09	make	0.10	work	0.14	sleep	0.10	phone	0.11	audiobooks	0.11
9	like	0.09	listen	0.10	sleep	0.11	new	0.10	pass	0.11	audio_book	0.11
10	call	0.09	thing	0.10	reflect	0.11	long	0.09	new	0.10	work	0.10

8 a. Work

9 The network for “work” and co-occurrence to other terms is shown in Figure 4. For the current
10 driving situation, individuals rarely noted any work-related tasks they were willing to do while
11 commuting. Often, people reported a preference to release stress rather than work. In AV, the
12 tasks related to work are more clearly identified: people reported interest in planning for their day
13 (make to-do lists, prepare, get ready for the day), read emails or news, and do some general work

1 (get things done). The significance of the co-occurrences are high, which demonstrate a level of
2 statistical reliability in the findings.

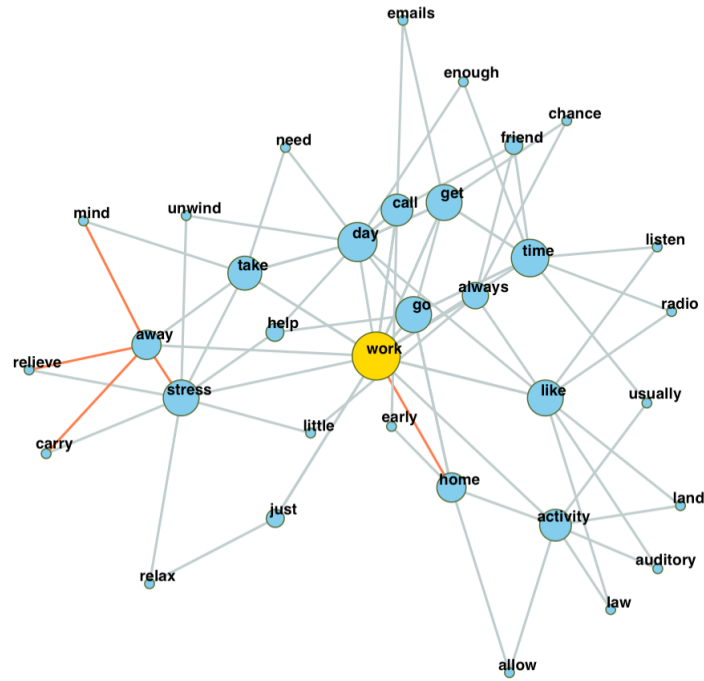
3 *b. Relax*

4 The network for “relax” and co-occurrence to other terms is shown in Figure 5. In current driving
5 situations, people would like to relax when they drive home by ‘unwinding their minds’, ‘reflecting
6 on the day’, sleeping, or listening to music or podcasts. When using an AV, people wanted to
7 similarly relax. However, they noted a greater interest in reading (personal emails, news, social
8 media, books) when compared to current driving. In AV mode, they also noted a preference to
9 listening to books, in addition to music and podcasts.

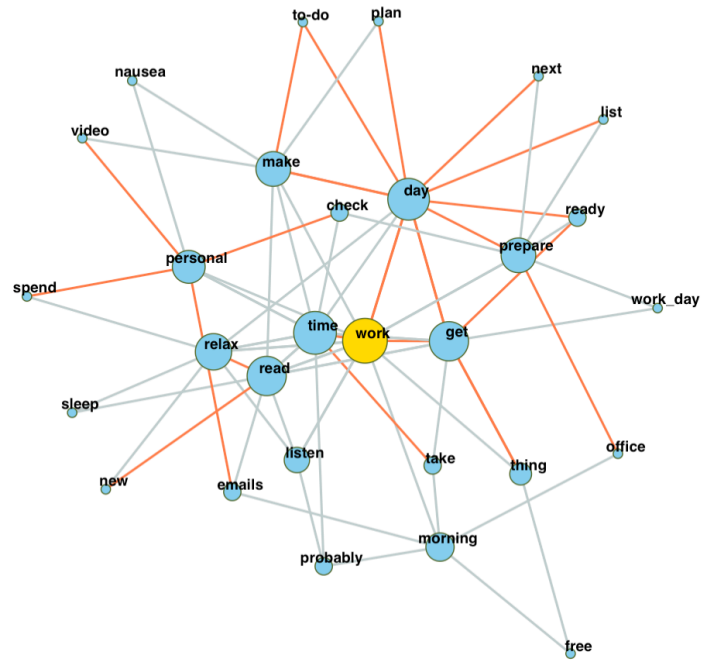
10 The critical difference is that tasks wanted when using AVs require more attention than using
11 current vehicle systems. The type of distraction changed from auditory (listen) to visual (read)—the
12 things listening to also changed from music or podcast to books.

13 *c. Listen*

14 The network for “listen” and co-occurrence to other terms is shown in Figure 6. For both the current
15 and AV situation, the most popular things to listen to were “radio,” “music,” “podcast.” The top
16 two activities have a high significance value (Current: radio=0.58, music=0.47; AV: music=0.53,
17 radio=0.42). This indicates high statistical reliability with “listen.” Similar to what was observed
18 in “relax”, most individuals were interested in listening to audio books in AV mode. Listening to
19 audio books require listeners to follow a story, imagine characters, and anticipate what will happen
20 next; this often requires greater attention than listening to music or podcast (Nowosielski et al.,
21 2018).

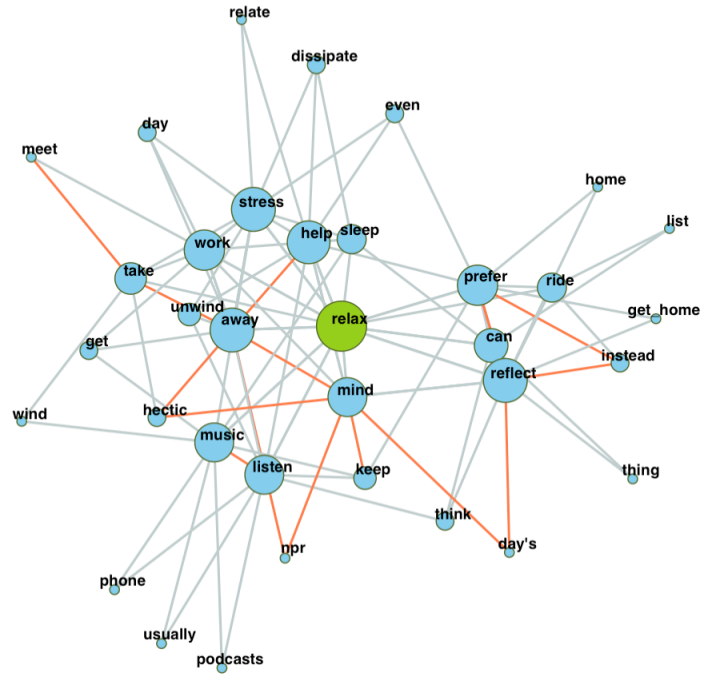


(a) Current situation

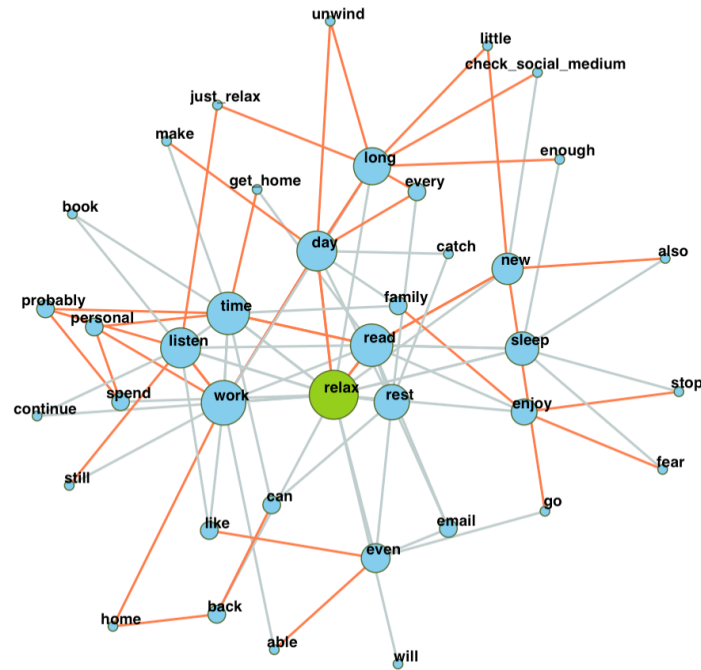


(b) Using an AV

Figure 4. Co-occurrence network for the target term "work" for the current driving situations (a) and for the situation with AV (b). The number of edges grew to 13 (current) and 11 (AV) from each node. Note: Edges with Dice statistics > 50% are shown in orange.

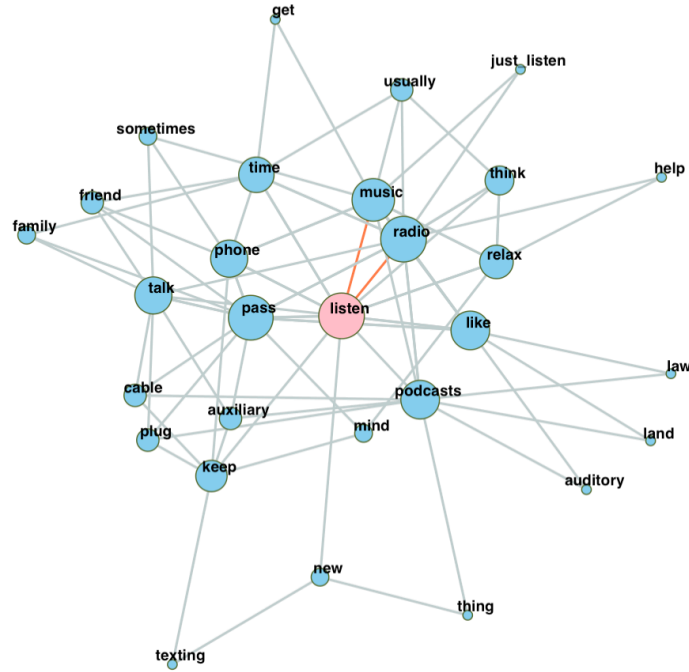


(a) Current situation

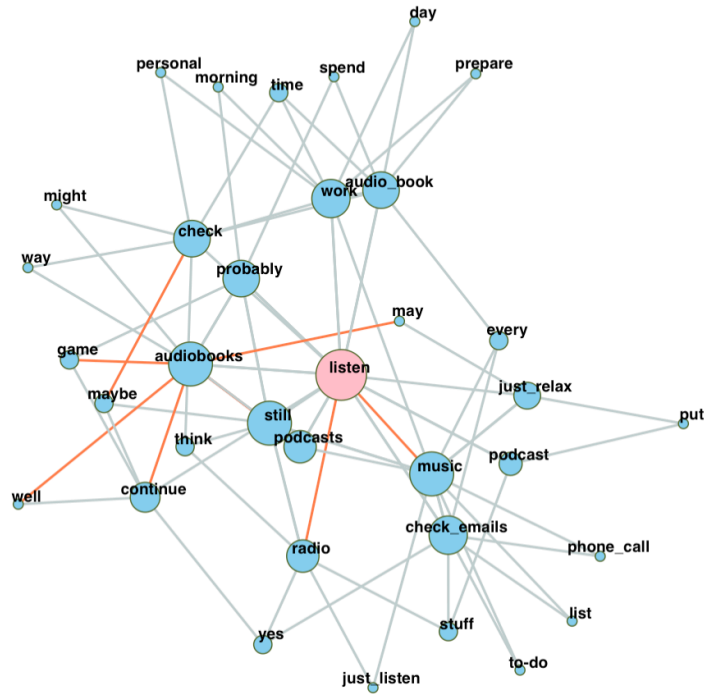


(b) Using an AV

Figure 5. Co-occurrence network for target word "Relax". (a) shows the result in current situation and (b) shows for using an AV. The limit for number of edges grew from each node was 13 and 14 correspondingly. Note: Edges with Dice statistics > 50% are shown in orange.



(a) Current situation



(b) Using an AV

Figure 6. Co-occurrence network for target term "Listen." (a) shows the result in the current situation and (b) shows for using an AV. The limit for the number of edges grew from each node was 12 and 13 correspondingly. Note: Edges with Dice statistics > 50% are shown in orange.

1 **4.3 Driver Perception Changes with AVs**

2 Changes in drivers’ perception with AV were examined using text clustering and participant clus-
3 tering. Text clustering grouped the responses while participant clustering grouped the participants
4 by their responses.

5 *4.3.1 Text Clustering*

6 Responses were examined in cluster groups from 2 to 10. With each additional cluster, a new dom-
7 inant term is observed and the goal is to assess whether the new groupings add value. Specifically,
8 for our data, when the number of clusters is 2, “work” is the dominant term in one group and “lis-
9 ten” is the dominant words in the second group. When the number of clusters is 3, we still have
10 “work” and “listen” as the dominant term in their respective groups, but “relax” is a dominant term
11 in a third group. When the number of clusters was set to 10, the cluster “Cl_listen” was separated
12 into two groups with the top term ”listen”. This was not meaningful for the purpose of this study,
13 so we stopped at nine clusters. The three clusters that included “nothing, none, no” were then man-
14 ually grouped into one cluster as these terms have similar meanings. This resulted in 7 unique text
15 clusters. Names were then assigned to clusters according to their top frequent terms. “Cl_”s were
16 attached to the names to distinguish between terms and clusters.

Table 3. Top five terms for each of the seven clusters

Cluster name	Components #	Top terms	Corresponding frequency
1. Cl_Work-related	2100	work, time, listen, get, phone	264, 199, 158, 133, 108
2. Cl_Listen	113	listen, music, radio, just_listen, relax	84, 62, 45, 22, 17
3. Cl_Relax	43	relax, help	43, 6
4. Cl_Sleep	29	sleep, relax, read, time	29, 6, 2, 1
5. Cl_Read	26	read, relax, ect, time	26, 5, 1, 1
6. Cl_Good	91	good	91
7. Cl_Nothing	168	none, no, nothing	94, 47, 43

1 The cluster *Cl_work-related* mainly gathered responses that showed a willingness to work while
2 commuting. As observed in Table 3, it also contains a mix of non-driving tasks (time, listen, get,
3 phone) that were all work-related. For example, “time” was to plan a schedule, time to work, or
4 extra time to work. The term “relax” was observed in four text clusters (listen, relax, sleep, read),
5 but the term’s association was different for each one. For example, some participants would note
6 that they wanted to “relax by reading” or “sleep to relax”. These text clusters were then used for
7 grouping participants in the next section.

8 4.3.2 Participant Clustering

9 The Silhouette analysis (Figure 7) showed that 8 or 9 clusters would provide the highest Silhouette
10 width. Given that it is best to select the smaller number of clusters possible, 8 is selected. The
11 number of participants in each cluster and the corresponding text cluster responses are shown in
12 Table 4.

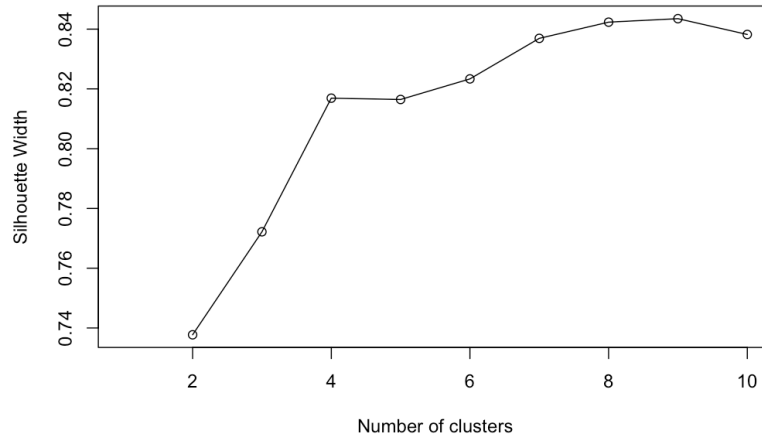


Figure 7. Silhouette width by number of clusters for participant clustering to help determine the optimal number of clusters.

Table 4. Eight clusters of participants.

#	Cl_ +	Cur mor	Cur eve	AV mor	AV eve	#	Cl_ +	Cur mor	Cur eve	AV mor	AV eve
1 (436)	Work-related	436	436	427	410	5 (9)	Work-related			9	7
	Listen				3		Listen				
	Relax				5		Relax				2
	Sleep			3	4		Sleep				
	Read						Read				
	Good			4	8		Good				
	Nothing			2	6		Nothing	9	9		
2 (28)	Work-related			1		6 (24)	Work-related			7	13
	Listen			1			Listen				
	Relax				1		Relax				1
	Sleep			1			Sleep				
	Read						Read				
	Good			4	8		Good	24	24	17	10
	Nothing	28	28	25	27		Nothing				
3 (31)	Work-related			30	30	7 (21)	Work-related	19	19	1	2
	Listen	31	31				Listen			2	
	Relax				1		Relax			3	4
	Sleep			1			Sleep			3	4
	Read						Read	2	2	8	7
	Good						Good			1	1
	Nothing						Nothing			3	3
4 (15)	Work-related	2	2			8 (10)	Work-related			7	6
	Listen	13	13	10	10		Listen			1	1
	Relax			2	2		Relax	8	8	1	1
	Sleep			1	1		Sleep	1	1	1	1
	Read			2	2		Read	1	1		
	Good						Good				
	Nothing						Nothing				1

Notes: *Cl_ +* refers to the name of the clusters from the text clustering results, which is consistent with “Cluster name” in Table 3. Columns *Cur mor*, *Cur eve*, *AV mor*, *AV eve* correspond to the four conditions analyzed. Numbers in parentheses are the size of the cluster.

Participant in clusters 1, 2, and 4 tend to maintain their behavior regardless of how they commute. About 76% of all valid participant responses (n=574) will always perform work or work-related tasks while commuting; about 4.9% and 2.6% of drivers will do nothing or listen to music/radio/podcast respectively, regardless of the vehicle systems. Participants in cluster 6 tend to have an overall positive attitude toward doing non-driving tasks during commuting (4.2%). One-third of them would work or want to involve in other mixed secondary tasks while commuting with AV.

1 Participants in clusters 3, 5, 7, and 8 reported changing their behavior when they are able to
2 commute using AV when compared to current driving. Participant cluster 3 tend to change from
3 listening (5.4%) to work; participant cluster 5 changed from nothing to (1.6%) to work; participant
4 cluster 7 changed from work to various behaviors (3.7 %); participant cluster 8 changed from relax
5 to work (1.7 %).

6 **5 DISCUSSION AND CONCLUSION**

7 In this study, we examined and compared drivers' preferences for engaging in non-driving tasks
8 while commuting in current vehicle systems and in future AVs. Text data was collected from an
9 open-ended survey and converted to document-to-term matrices for analysis.

10 Our first research question asked, "What changes are expected in the drivers' general attitude
11 towards performing non-driving tasks while commuting?". This was examined using sentiment
12 analysis. Our findings showed an overall positive attitude toward conducting non-driving tasks
13 while commuting in future AVs when compared to current vehicle systems.

14 The results show that participants, who are knowledge workers, preferred to use their commute
15 time for non-driving tasks in order to be productive. Being "productive" includes performing work
16 tasks to directly improve productivity; doing transition tasks (e.g., scheduling, planning) to better
17 transition mentally and physically between home and work; and even performing non-work tasks
18 (e.g., listening to music, sleeping) to be better energized for their next destination (Jachimowicz
19 et al., 2021).

20 In terms of the sentiment score for different conditions, commuting with AV in the morning
21 had the highest mean sentiment score, i.e., most positive. The findings show that drivers are more
22 positive toward conducting non-driving tasks in future AVs when compared to the current situation,
23 especially in the morning when they need to transition from home to work. During the morning

1 commute, the driver may prefer to conduct work-related tasks, which cannot be satisfactorily ac-
2 complished with current vehicle systems.

3 Our second research question asked, “What changes are expected between non-driving tasks
4 that are non-work and work-related?” This was examined using term frequency and term co-occurrence
5 analysis methods. Compared to current vehicle systems, people were less willing to listen to mu-
6 sic/radio/ podcasts and more willing to work with future AVs. Results that support this trend include
7 the lower number of “listen” and higher number of “work” for morning and evening commute with
8 future AVs. People expressed interest in working on more attention demanding tasks while riding
9 in future AVs. This was identified as we delved deeper into what people want to work on or how
10 they want to relax in the vehicle. While using an AV, visually demanding tasks such as reading
11 emails, making to-do lists, and preparing for their day co-occurred with “work” with high signifi-
12 cance. That is, the distracting tasks changed from auditory to visual. For current driving situations,
13 there did not appear to be a consistent answer for how people preferred to work in a vehicle. As for
14 relaxation, ‘sleep’ and ‘read’, were detected as the top-ranking words in a cluster for AV morning.
15 People preferred extra sleep when they drive to work in the morning.

16 In the future, people will most likely seek larger screens and keyboards or new interfaces for
17 non-driving tasks, since they are more willing to check emails, make to-do lists, and other prepa-
18 rations for the work day. The small screens on phones will not meet the need. This is supported
19 by the co-occurrence analysis for “work” and the drop in the frequency of the word “phone”. New
20 interfaces like see-through Augmented Reality devices or projection on the windshield may also
21 help users keep their eyes on the road while providing a better immersing experience (Rusch et al.,
22 2013).

23 Our third research question asked, “How will drivers’ perception towards performing non-
24 driving tasks change after using AVs?”. This was examined using text clustering and participant
25 clustering. Our findings show that most drivers would maintain their habits in the vehicle while
26 commuting (83.4%): 76.0% participants will keep conducting work-related tasks; 4.9% will keep

1 their listening habits; 2.6 % would still not perform any non-driving tasks. Preference for non-
2 driving tasks seems to be a personal habit. Designers may want to consider personalized designs in
3 future AVs to meet demands and support individual preferences for non-driving task engagement.

4 Compared with current driving situations, 8.7% of participants (from three participant clusters)
5 would abandon their previous tasks (“listen,” “relax,” and “do nothing”) to do work or work-related
6 tasks in the future if they are able to use AVs for commuting. However, only 3.7% of participants
7 (from one participant cluster) would act in the opposite direction once they are able to commute with
8 AVs; these individuals reported being distracted by other tasks and willingness to abandon work-re-
9 lated tasks. In terms of application, future designs should consider these differences in developing
10 assistance systems, and adapt appropriately given drivers willingness to change their behavior in
11 AVs.

12 In terms of study limitations, a term frequency analysis was used at the start and this may pose
13 some issues with polysemy (using similar terms in different contexts) and synonymy (using dif-
14 ferent terms while meaning the same thing). There are also target terms (e.g., “listen”, “work”,
15 “relax”) that may be difficult for a computer algorithm to distinguish. While there are no perfect
16 solutions, we based our existing literature in text mining, which shows that many text analysis use
17 term frequency as a first step for understanding patterns from unstructured text data. (Das et al.,
18 2019).

19 Our findings show that most people have a positive attitude towards automated vehicles, which
20 they perceive as providing opportunities to engage in non-driving tasks during their commute. To
21 provide a safer and more productive environment, the in-vehicle interfaces should be designed
22 to enable effective task switching between more visually-involved work-related/relaxing tasks and
23 driving. Future studies can also consider the overall experience with non-driving tasks and consider
24 ways to better tailor the tasks toward the human operator.

Acknowledgement

This research is funded by National Science Foundation (Grant No. NSF 1840031). The findings and opinions expressed in this article are those of the authors and do not necessarily reflect those of NSF.

References

- Alt, F., Kern, D., Schulte, F., Pfleging, B., Shirazi, A. S., and Schmidt, A. (2010). Enabling micro-entertainment in vehicles based on context information. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '10, page 117–124, New York, NY, USA. Association for Computing Machinery.
- Bálint, A., Flannagan, C. A., Leslie, A., Klauer, S., Guo, F., and Dozza, M. (2020). Multitasking additional-to-driving: Prevalence, structure, and associated risk in SHRP2 naturalistic driving data. *Accident Analysis and Prevention*, 137(August 2019):105455.
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., and Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30):774.
- Cyganski, R., Fraedrich, E., and Lenz, B. (2015). Travel-Time Valuation for Automated Driving: A Use-case-driven study. In *Proceedings of the 94th Annual Meeting of the Transportation Research Board*, Washington DC, United States. Transportation Research Board.
- Das, S., Dutta, A., Lindheimer, T., Jalayer, M., and Elgart, Z. (2019). YouTube as a Source of Information in Understanding Autonomous Vehicle Consumers: Natural Language Processing Study. *Transportation Research Record*, 2673(8):242–253.
- Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology*.
- Drucker, P. F. (1999). Knowledge-worker productivity: The biggest challenge. *California management review*, 41(2):79–94.
- Feinerer, I. (2015). Introduction to the tm Package: Text Mining in R. *R vignette*.
- Frick, D. E. (2010). Motivating the knowledge worker. Technical report, DEFENSE INTELLIGENCE AGENCY WASHINGTON DC.

- 1 Ghazizadeh, M., McDonald, A. D., and Lee, J. D. (2014). Text mining to decipher free-response
2 consumer complaints: Insights from the NHTSA vehicle owner's complaint database. *Human*
3 *Factors*, 56(6):1189–1203.
- 4 Hudson, J., Orviska, M., and Hunady, J. (2019). People's attitudes to autonomous vehicles. *Trans-*
5 *portation Research Part A: Policy and Practice*, 121(August 2018):164–176.
- 6 Hulse, L. M., Xie, H., and Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships
7 with road users, risk, gender and age. *Safety Science*, 102(October 2017):1–13.
- 8 Humagain, P. and Singleton, P. A. (2020). Would you rather teleport or spend some time commut-
9 ing? investigating individuals' teleportation preferences. *Transportation research part F: traffic*
10 *psychology and behaviour*, 74:458–470.
- 11 Jachimowicz, J. M., Cunningham, J. L., Staats, B. R., Gino, F., and Menges, J. I. (2021). Be-
12 tween home and work: commuting as an opportunity for role transitions. *Organization Science*,
13 32(1):64–85.
- 14 Jing, L., Zhou, L., Ng, M. K., and Huang, J. Z. (2006). Ontology-based distance measure for text
15 clustering. In *Proc. of SIAM SDM workshop on text mining, Bethesda, Maryland, USA*.
- 16 Jones, K. S. (1972). A statistical interpretation of term specificity and its application in retrieval.
- 17 Kaufman, L. and Rousseeuw, P. (2008). *Partitioning Around Medoids (Program PAM)*. John Wiley
18 & Sons, Ltd.
- 19 Kolesnikova, O. (2016). Survey of Word co-occurrence measures for collocation detection. *Com-*
20 *putacion y Sistemas*, 20(3):327–344.
- 21 Lee, L. H., Wan, C. H., Rajkumar, R., and Isa, D. (2012). An enhanced support vector machine
22 classification framework by using euclidean distance function for text document categorization.
23 *Applied Intelligence*, 37(1):80–99.
- 24 Malokin, A., Circella, G., and Mokhtarian, P. L. (2019). How do activities conducted while com-
25 muting influence mode choice? Using revealed preference models to inform public transporta-
26 tion advantage and autonomous vehicle scenarios. *Transportation Research Part A: Policy and*
27 *Practice*, 124(December 2018):82–114.

- 1 McEvoy, S. P., Stevenson, M. R., and Woodward, M. (2006). The impact of driver distraction on
2 road safety: Results from a representative survey in two Australian states. *Injury Prevention*,
3 12(4):242–247.
- 4 Milakis, D., Van Arem, B., and Van Wee, B. (2017). Policy and society related implications of
5 automated driving: A review of literature and directions for future research. *Journal of Intelligent*
6 *Transportation Systems: Technology, Planning, and Operations*, 21(4):324–348.
- 7 Murtagh, F. and Legendre, P. (2014). Ward’s Hierarchical Agglomerative Clustering Method:
8 Which Algorithms Implement Ward’s Criterion? *Journal of Classification*.
- 9 Naldi, M. (2019). A review of sentiment computation methods with R packages. *CoRR*,
10 abs/1901.08319.
- 11 Nowosielski, R. J., Trick, L. M., and Toxopeus, R. (2018). Good distractions: Testing the effects
12 of listening to an audiobook on driving performance in simple and complex road environments.
13 *Accident Analysis & Prevention*, 111:202–209.
- 14 Penmetsa, P., Adanu, E. K., Wood, D., Wang, T., and Jones, S. L. (2019). Perceptions and ex-
15 pectations of autonomous vehicles – A snapshot of vulnerable road user opinion. *Technological*
16 *Forecasting and Social Change*, 143(October 2018):9–13.
- 17 Reja, U., Manfreda, K. L., Hlebec, V., and Vehovar, V. (2003). Open-ended vs. Close-ended
18 Questions in Web Questionnaires. *Developments in Applied Statistics*, 19(January):159–177.
- 19 Rinker, T. W. (2021). *sentimentr: Calculate Text Polarity Sentiment*. Buffalo, New York. version
20 2.9.0.
- 21 Ronen, F. and James, S. (2007). The text mining handbook: Advanced approaches in analyzing
22 unstructured data.
- 23 Rusch, M. L., Schall Jr, M. C., Gavin, P., Lee, J. D., Dawson, J. D., Vecera, S., and Rizzo, M.
24 (2013). Directing driver attention with augmented reality cues. *Transportation research part F:*
25 *traffic psychology and behaviour*, 16:127–137.
- 26 Schütze, H. and Silverstein, C. (1997). Projections for efficient document clustering. In *ACM*
27 *SIGIR Forum*, volume 31, pages 74–81. ACM New York, NY, USA.

- 1 Teodorovicz, T., Kun, A. L., Sadun, R., and Shaer, O. (2022). Multitasking while driving: A time
2 use study of commuting knowledge workers to assess current and future uses. *International*
3 *Journal of Human-Computer Studies*, 162:102789.
- 4 Teodorovicz, T., Sadun, R., Kun, A. L., and Shaer, O. (2021). *Working from Home During*
5 *COVID19: Evidence from Time-use Studies*. Harvard Business School.
- 6 Thomas, K. L., Peter, W. F., and Darrell, L. (1998). An introduction to latent semantic analysis.
7 *Discourse Processes*, 25:259–284.
- 8 Varghese, V. and Jana, A. (2018). Impact of ICT on multitasking during travel and the value of
9 travel time savings: Empirical evidences from Mumbai, India. *Travel Behaviour and Society*,
10 12:11–22.
- 11 Wiedemann, G. and Wiedemann (2016). *Text mining for qualitative data analysis in the social*
12 *sciences*, volume 1. Springer.
- 13 Young, S. D., Adelstein, B. D., and Ellis, S. R. (2007). Demand characteristics in assessing motion
14 sickness in a virtual environment: Or does taking a motion sickness questionnaire make you
15 sick? In *IEEE Transactions on Visualization and Computer Graphics*.