Transportation Research Part F: Psychology and Behaviour 94 (2023) 305–320

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

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

Yilun Xing [a], Linda Ng Boyle [a,*], Rafaella Sadun [b], John D. Lee [c], Orit Shaer [d], Andrew Kun [e]

[a] University of Washington, Seattle, WA, United States of America
[b] Harvard University, Boston, MA, United States of America
[c] University of Wisconsin-Madison, Madison, WI, United States of America
[d] Wellesley College, Wellesley, MA, United States of America
[e] University of New Hampshire, Durham, NH, United States of America

ARTICLE INFO

Keywords:
Automated vehicles
Text mining
Commuting
Non-driving tasks

ABSTRACT

Automated vehicles (AVs) offer human operators the opportunity to participate in non-driving activities while on the move. In this study, we examined and compared drivers' perception of non-driving activities in two driving modes: highly AVs in the future and current vehicle systems, where the human operator is still responsible for controlling the vehicle such as braking and steering. The study used a survey distributed through an online paid marketplace platform called Lucid, which included open-ended questions soliciting participants' perceptions of non-driving activities given a work commute scenario for each driving mode. Text mining and clustering analysis were used to analyze the responses of 752 participants to four open-ended survey questions. Results showed that drivers had a more positive sentiment towards future automated vehicles compared to current systems. The most reported non-driving activities overall were "work", "listen", and "relax"; were "listen" for current vehicle systems and "work" for AVs. The study also captured the changes in drivers' perception from current systems to AV systems. The findings indicated that most drivers (83.4%) would continue their current non-driving activities, with 76.0% continuing to perform work or work-related activities. Approximately 8.7% of respondents would switch from their current tasks to work-related tasks in an AV, while 3.7% would do the opposite—abandon work-related tasks to do other activities. The study suggests that working while commuting will be an advantage of AVs, highlighting the need to understand how people can work productively as we move forward with automated vehicles.

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%),

* Corresponding author.
E-mail address: linda@uw.edu (L.N. Boyle).

https://doi.org/10.1016/j.trf.2023.01.015
Received 4 November 2021; Received in revised form 3 November 2022; Accepted 18 January 2023

Available online 8 March 2023

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 in 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 worker's overall well-being. To effectively provide drivers the opportunity to perform non-driving tasks in AVs, we need to first understand people's perceptions 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 their preference or ability to engage in other non-driving tasks (Hudson et al., 2019, Milakis et al., 2017, Hulse et al., 2018, Penmetsa et al., 2019). In one study conducted by Cyganski et al. (2015) on potential activities in AV, music and talking to passengers were reported to be the most often conducted activities while driving. Their study showed that working while traveling played a minor role and was rarely considered an advantage of AVs by participants.

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

1. What changes in the drivers' general attitude is expected when performing non-driving tasks during commute times (morning, evening) in each system mode (AV, manual)?
2. What changes are expected between non-driving tasks that are non-work and work-related?
3. Are there comparative differences in how drivers expect to perform non-driving tasks in a future world that includes commuting with AVs?

2. Data source

A time-use survey was distributed using the online platform Lucid, which partners with several companies to recruit individuals to answer online surveys (a copy of the survey is available in the paper of 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 different scenarios (morning commute, and 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?

These 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.

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

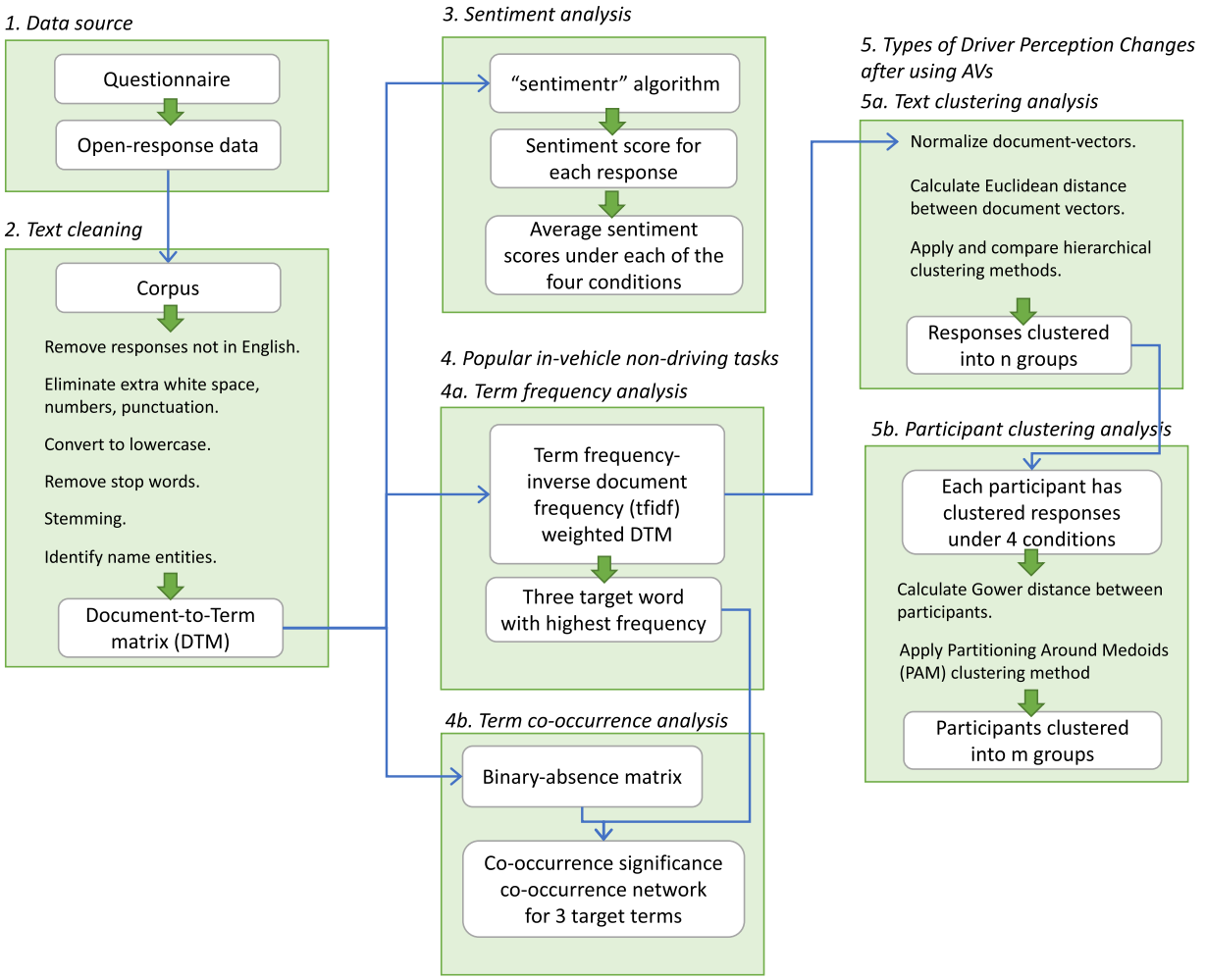


Fig. 1. Flow chart of analysis procedure.

There were 756 commuting knowledge workers that responded to the survey. Of these, four were removed given incomplete information for a total of 752 used for the subsequent analysis. Among the 752 participants, 346 are female and 406 are male. Participants' age ranges from 18 to 79 with an average of 38.02. Most participants (96.5%) reside in the US with 21 NA's and 5 from other counties such as France and India (the country of residence were mapped with lat/long coordinates).

3. Methods

A text mining approach was used to examine the open-ended responses. There are two concepts commonly used in text mining: *document* is a unit of analysis of textual data, which corresponds to each response in a question (Feldman & Sanger, 2007), and *terms* are the components that make up the *document* (Ghazizadeh et al., 2014).

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

The open-ended responses provide insights into drivers' non-driving tasks. We, therefore, examined the changes in perception toward non-driving tasks if AVs were available (e.g., listening to music during the morning commute in manual driving, to work during the morning commute in AVs) (Box 5). All responses were clustered into different term groups and each group was labeled with the most frequent term in that group (Box 5a). Participants' responses were converted from text data to categorical data. Because each participant had four responses (or four commute conditions), we have four independent variables for each participant, and these are used for the participant clustering analysis (Box 5b). The clustering results show how each group's perception of non-driving task engagement changed from the current condition to an envisioned future when they will be able to use AVs.

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 the 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, and punctuation and convert to lowercase.
   Our analysis focused on term frequency, and 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,” and “commute”.
5. Stemming.
   Stemming is a method that converts words to their radicals, e.g., “working” “worked” and “works” becomes “work”.
6. Identify named entities.
   Some words appear together frequently and represent a single concept. This includes phrases such as “social media” and “phone call”. Words that appeared together ten or more times in the dataset were combined into a single term with “_”.
7. Transform to Document-to-Term matrix (DTM).
   Corpus data were transformed to a DTM and then used for analysis. Each row of the DTM refers to a document, and each column of the DTM refers to a term. Each entry of the DTM represents the number of occurrences of a term in a document.

After the text cleaning process, the number of unique terms in our corpus for the Current Morning survey item was reduced from 854 terms to 632; Current Evening from 756 to 564; AV Morning from 928 to 703; AV Evening from 847 to 641. The number of non-missing values for the Current Morning and Current Evening survey responses was reduced to 673; AV Morning to 696; AV Evening to 610 (e.g., responses such as “no”, and “0” were removed through the processes). Therefore, for the text clustering procedure, there were (696 + 696+ 673+ 610 =) 2675 responses included. For the participant clustering procedure, since a response for each of the four conditions is required, 574 participants’ data were used as input for the clustering model. Among the 574 participants, 249 are female and 325 are male. Participants’ age ranges from 21 to 79 with an average of 38.0. Most participants (97.0%) reside in the US with 14 NA’s and 3 from other counties such as India (the country of residence was 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 its 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 words in a sentence. Whether the negators are considered depends on whether the lexicon includes negators or whether the user searches for the negators and reverses 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 valence shifters that will have an impact on the sentiment of the polarized word. The words in this subset are tagged as neutral, negator, amplifier, or de-amplifier.

Each neutral holds no value, while each valance shifter is assigned a weight varied by the valence shifter type. Amplifiers become de-amplifiers if the subset contains an odd number of negators. A detailed description of how the weights are assigned can be seen in the package help file. The accumulated weight is then added to the sentiment score of the polarized word. Last, these weighted polarized context subset scores are summed and divided by the squared root of the word count yielding an unbounded sentiment score for each sentence. The average sentiment score of each questionnaire response is to get the mean of all the sentences within it. A positive sentiment score indicates an overall positive attitude and the greater the value the more positive attitude the participant holds.

3.3. Term frequency analysis

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

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

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

However, this is not usually the case in practice. The most frequent terms are not necessarily the most informative ones; terms that appear frequently in a small number of documents but rarely in other documents tend to be more relevant and specific for that particular group of documents, and 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 the document frequency $df(t_j)$ represents the number of documents in which term t_j appears. In this study, *tfidf* was used for all analyses except the sentiment analysis, which uses the exact number of positive/negative terms to calculate the proportion of each type of term.

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 a 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 & Wiedemann, 2016, Kolesnikova, 2016). The formula used for calculating the Dice statistic of all responses A containing one term a and all responses B containing one term b is:

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

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

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

3.5. Text clustering

The term frequency method (*tfidf*) is shown to be as good as the semantic-based methods (e.g., Latent Semantic Indexing) for text clustering (Schütze & Silverstein, 1997). *tfidf* was chosen because it is more intuitive to tag each text cluster with their most

frequent term first (e.g., work) and then group participants based on their responses from the text cluster (goal of this study). The semantic method, which shares similar ideas with Principle Component Analysis (PCA), represents terms as a linear combination of other terms (Thomas et al., 1998). The findings are not as intuitive as term frequency methods, making it more difficult to tag the clustered answers to a term.

Each row of the *tfidf* weighted DTM matrix is a vector representation of a document. Distance between the two vectors was then measured to capture the similarity of the two documents and clustered into groups based on this similarity. Euclidean distance was used as it is widely used 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:

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

where $\widehat{DTM}_{i_1}$ and $\widehat{DTM}_{i_1}$ 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 & 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. The 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 the morning with AV) based on the Gower distance (Kaufman & 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 include 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. The silhouette method was applied to get the optimal number of clusters by examining the width of the clusters.

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

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

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 Fig. 2. There were more responses with positive sentiment scores than negative scores in all conditions, indicating an overall positive attitude towards doing non-driving tasks during commuting. Participants had a more positive sentiment toward doing non-driving tasks with AVs (M = 0.138, SD = 0.290) than in the current driving condition (M = 0.133, SD = 0.300); and more positive sentiment toward doing non-driving tasks during morning commutes (M = 0.145, SD = 0.299) than in the evening (M = 0.125, SD = 0.291); No significant differences were detected at $\alpha = 0.05$ using pair-wise t-tests.

When comparing the sentiment among the four conditions, results showed the highest positive sentiment for AV in the morning (M = 0.154, SD = 0.296) followed by current driving in the morning (M = 0.136, SD = 0.302), current driving in the evening

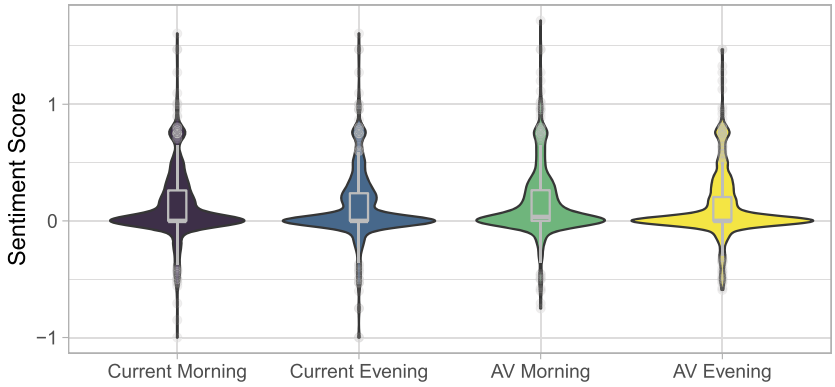


Fig. 2. Violin plot for the sentiment scores for the four driving conditions.

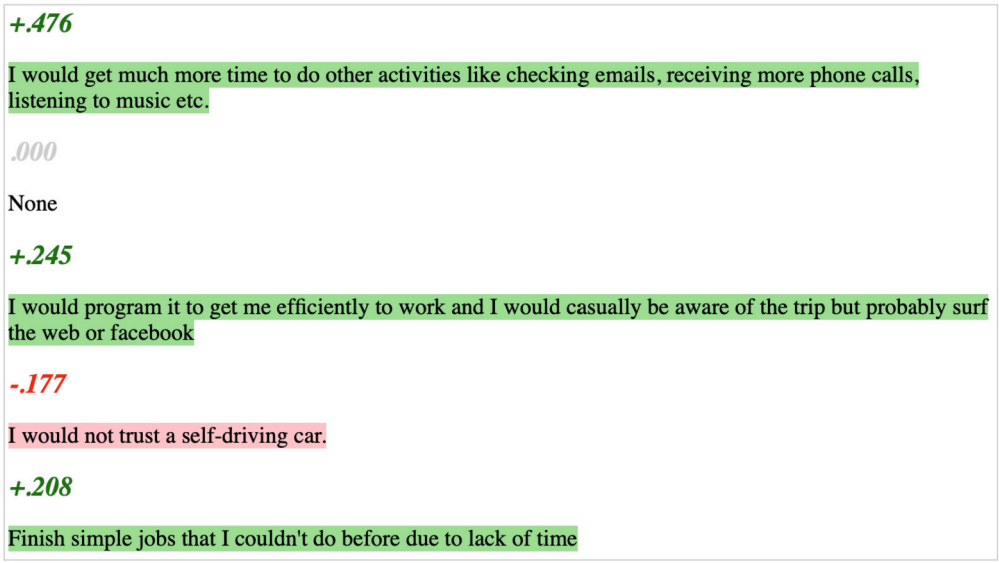


Fig. 3. Examples of responses highlighted with green and red to show their polarity. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

(M = 0.129, SD = 0.299), and lastly, AV in the evening (M = 0.122, SD = 0.283). A pairwise t-test did not show any significant differences ($p > 0.05$). Though the two AV conditions did not both have more positive scores than the two current driving conditions, there were fewer responses with more negative sentiment scores, i.e., few responses with a sentiment score below −0.5. Examples of sentiment scores are shown in Fig. 3.

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 the morning and evening. People reported a preference for work-related tasks in the morning and a preference for relaxing or listening to the 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.

	Current Morning		Current Evening		AV Morning		AV Evening	
Rank	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 shows 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 together to visualize their change in frequency. The frequency of "work" and "relax" will exceed "listen" when AV is used for commuting.

4.2.2. Term co-occurrence analysis

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

a. Work

The network for "work" and co-occurrence to other terms is shown in Fig. 4. For the current driving situation, individuals rarely noted any work-related tasks they were willing to do while commuting. Often, people reported a preference to release stress rather than work. In AV, the tasks related to work are more clearly identified: people reported interest in planning for their day (making to-do lists, preparing, getting ready for the day), reading emails or news, and doing some general work (getting things done). The significance of the co-occurrences is high, which demonstrates a level of statistical reliability in the findings.

b. Relax

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

The critical difference is the network of tasks that participants reported interest in performing when using AVs. For example, the most frequent listening tasks changed from listening to music (in current systems) to listening to audiobooks (in AVs). This larger

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

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

network of listening tasks may require more attention, which is the reason they may be preferred in AVs. Preference was also given toward visual (reading) tasks when compared to auditory (listening) tasks in AVs.

c. Listen

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

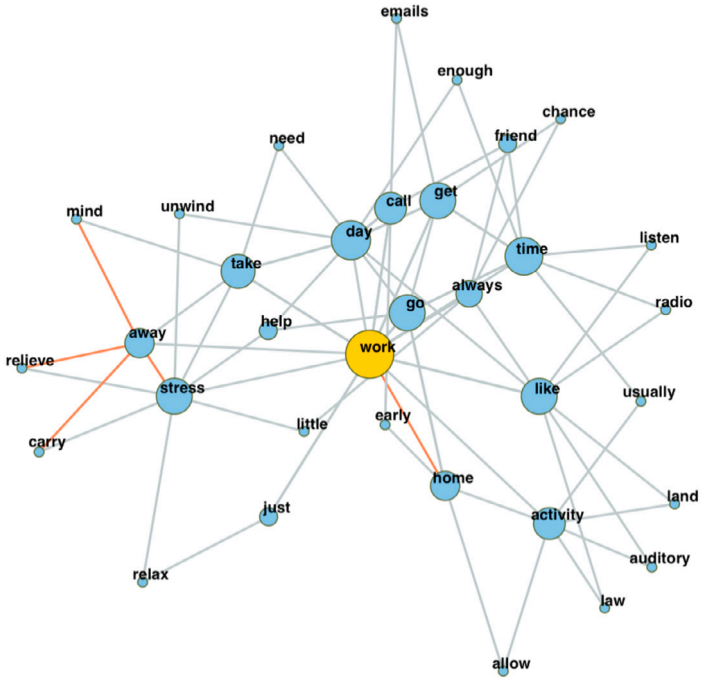
4.3. Driver perception changes with AVs

Changes in drivers' perception of AV were examined using text clustering and participant clustering. Text clustering grouped the responses while participant clustering grouped the participants by their responses.

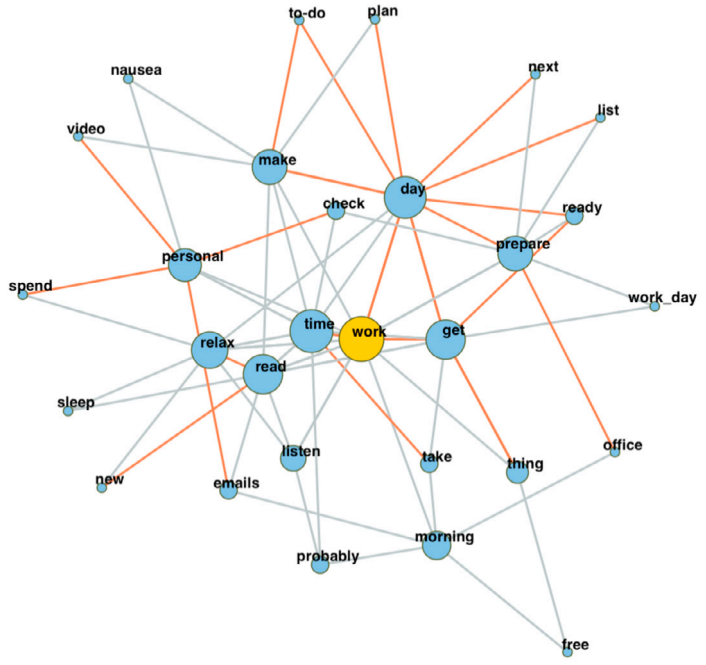
4.3.1. Text clustering

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

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



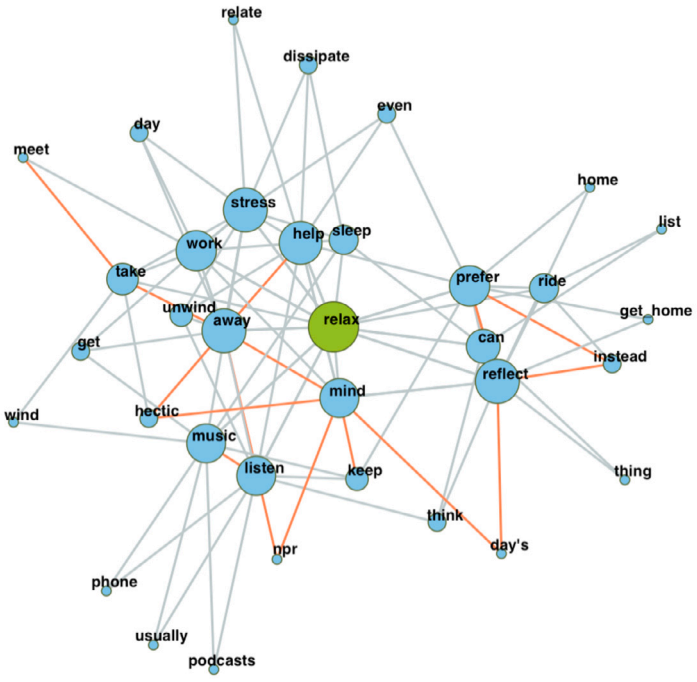
(a) Current situation



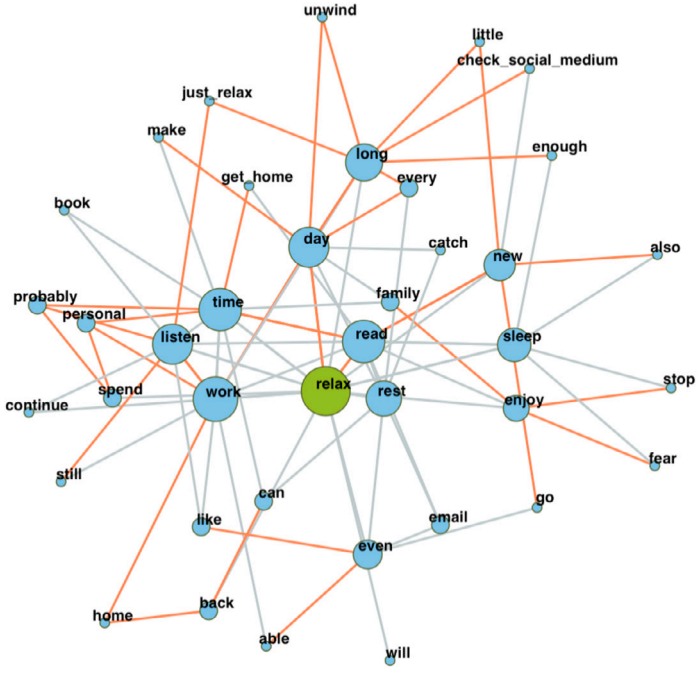
(b) Using an AV

Notes: The limit for the number of edges that grew from each node was 13 (current) and 11 (AV). Edges with Dice statistics > 50% are shown in orange. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Fig. 4. Co-occurrence network for the target term "work" for (a) the current driving situations and (b) for future situations using AV.



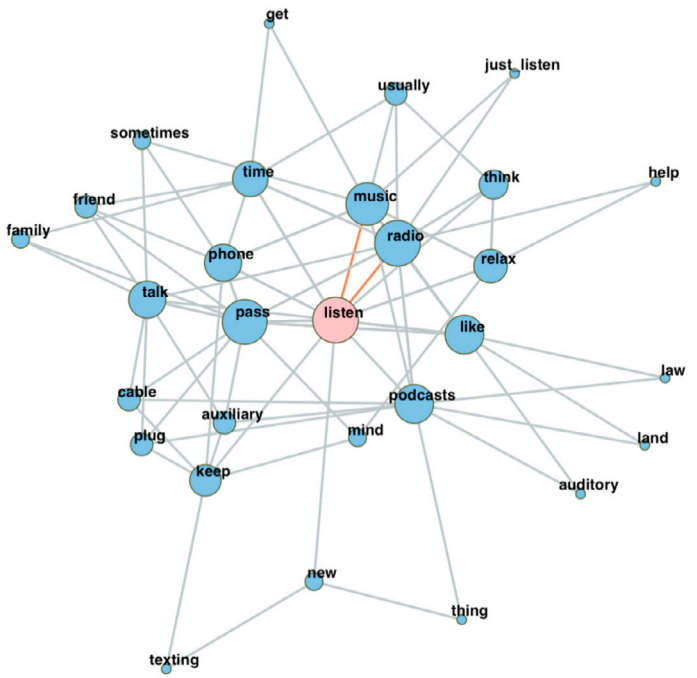
(a) Current situation



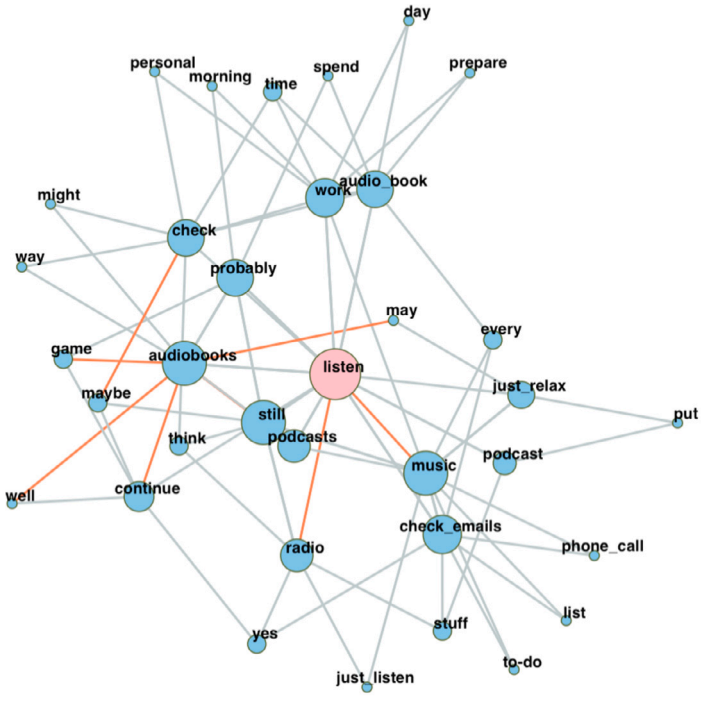
(b) Using an AV

Notes: The limit for the number of edges that grew from each node was 13 (current) and 14 (AV). Edges with Dice statistics > 50% are shown in orange. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Fig. 5. Co-occurrence network for target term "relax" for (a) the current driving situation and (b) for future situations using AV.



(a) Current situation



(b) Using an AV

Notes: The limit for the number of edges that grew from each node was 12 (current) and 13 (AV). Edges with Dice statistics > 50% are shown in orange. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Fig. 6. Co-occurrence network for target term "Listen." (a) shows the result in the current situation and (b) shows for using an AV.

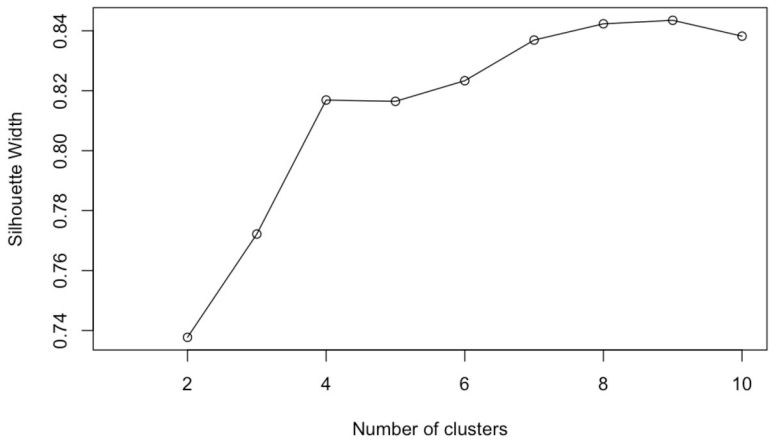


Fig. 7. Silhouette width by number of clusters.

Table 4
Eight clusters of participants.

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

4.3.2. Participant clustering

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

Participants 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.

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

5. Discussion and conclusion

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

Our first research question asked, "What changes in the drivers' general attitude is expected when performing non-driving tasks during commute times (morning, evening) in each system mode (AV, manual)?". This was examined using sentiment analysis. Our findings showed an overall positive attitude toward conducting non-driving tasks while commuting in future AVs when compared to current vehicle systems.

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

In terms of the sentiment score for different conditions, commuting with AV in the morning had the highest mean sentiment score, i.e., the most positive. The findings show that drivers are more positive toward conducting non-driving tasks in future AVs when compared to the current situation, especially in the morning when they need to transition from home to work. During the morning commute, the driver may prefer to conduct work-related tasks, which cannot be satisfactorily accomplished with current vehicle systems.

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

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

Our third research question asked, "Are there comparative differences in how drivers expect to perform non-driving tasks in a future world that includes commuting with AVs?". This was examined using text clustering and participant clustering. Our findings show that most drivers would maintain their habits in the vehicle while commuting (83.4%): 76.0% participants will keep conducting work-related tasks; 4.9% will keep their listening habits; 2.6% would still not perform any non-driving tasks. Preference for non-driving tasks seems to be a personal habit. Designers may want to consider personalized designs in future AVs to meet demands and support individual preferences for non-driving task engagement.

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

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

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

CRediT authorship contribution statement

Yilun Xing: Conceptualization, Methodology, Validation, Formal analysis, Writing – Original Draft, Review and Editing, Visualization. **Linda Ng Boyle:** Methodology, Validation, Resources, Writing – Original Draft, Review and Editing, Supervision, Project Administration. **Rafaella Sadun:** Conceptualization, Investigation, Data Curation, Review and Editing. **John D. Lee:** Methodology, Writing - Original Draft, Review and Editing. **Orit Shaer:** Investigation, Review and Editing. **Andrew Kun:** Conceptualization, Review and Editing, Project Administration, Funding acquisition.

Data availability

The data that has been used is confidential.

Acknowledgement

This research is funded by National Science Foundation (Grant No. NSF 1839666; 1840085; 1840031; 1839484; 1839870). 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., & 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* (pp. 117–124). New York, NY, USA: Association for Computing Machinery.

Bálint, A., Flannagan, C. A., Leslie, A., Klauer, S., Guo, F., & Dozza, M. (2020). Multitasking additional-to-driving: Prevalence, structure, and associated risk in SHRP2 naturalistic driving data. *Accident Analysis and Prevention*, *137*(August 2019), Article 105455.

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *The Journal of Open Source Software*, *3*(30), 774.

Cyganski, R., Fraedrich, E., & 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., & 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*, *26*(3), 297–302.

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*. https://cran.r-project.org/web/packages/tm/vignettes/tm.pdf.

Feldman, R., & Sanger, J. (2007). The text mining handbook: Advanced approaches in analyzing unstructured data. Cambridge University Press.

Frick, D. E. (2010). Motivating the knowledge worker. Technical report, Defense Intelligence Agency Washington DC.

Ghazizadeh, M., McDonald, A. D., & Lee, J. D. (2014). Text mining to decipher free-response consumer complaints: Insights from the NHTSA vehicle owner's complaint database. *Human Factors*, *56*(6), 1189–1203.

Hudson, J., Orviska, M., & Hunady, J. (2019). People's attitudes to autonomous vehicles. *Transportation Research. Part A, Policy and Practice*, *121*(August 2018), 164–176.

Hulse, L. M., Xie, H., & Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science*, *102*(October 2017), 1–13.

Humagain, P., & Singleton, P. A. (2020). Would you rather teleport or spend some time commuting? Investigating individuals' teleportation preferences. *Transportation Research. Part F, Traffic Psychology and Behaviour*, *74*, 458–470.

Jachimowicz, J. M., Cunningham, J. L., Staats, B. R., Gino, F., & Menges, J. I. (2021). Between home and work: Commuting as an opportunity for role transitions. *Organization Science*, *32*(1), 64–85.

Jing, L., Zhou, L., Ng, M. K., & Huang, J. Z. (2006). Ontology-based distance measure for text clustering. In *Proc. of SIAM SDM workshop on text mining*. Maryland, USA: Bethesda.

Jones, K. S. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, *28*(1), 11–21.

Kaufman, L., & Rousseeuw, P. (2008). *Partitioning around medoids (program PAM)*. John Wiley & Sons, Ltd.

Kolesnikova, O. (2016). Survey of Word co-occurrence measures for collocation detection. *Computacion y Sistemas*, *20*(3), 327–344.

Lee, L. H., Wan, C. H., Rajkumar, R., & Isa, D. (2012). An enhanced support vector machine classification framework by using Euclidean distance function for text document categorization. *Applied Intelligence*, *37*(1), 80–99.

Malokin, A., Circella, G., & Mokhtarian, P. L. (2019). How do activities conducted while commuting influence mode choice? Using revealed preference models to inform public transportation advantage and autonomous vehicle scenarios. *Transportation Research. Part A, Policy and Practice*, *124*(December 2018), 82–114.

McEvoy, S. P., Stevenson, M. R., & Woodward, M. (2006). The impact of driver distraction on road safety: Results from a representative survey in two Australian states. *Injury Prevention*, *12*(4), 242–247.

Milakis, D., Van Arem, B., & Van Wee, B. (2017). Policy and society related implications of automated driving: A review of literature and directions for future research. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, *21*(4), 324–348.

Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification*, *31*, 274–295.

Naldi, M. (2019). A review of sentiment computation methods with R packages. CoRR. arXiv:1901.08319 [abs].

Nowosielski, R. J., Trick, L. M., & Toxopeus, R. (2018). Good distractions: Testing the effects of listening to an audiobook on driving performance in simple and complex road environments. *Accident Analysis and Prevention*, *111*, 202–209.

Penmetsa, P., Adanu, E. K., Wood, D., Wang, T., & Jones, S. L. (2019). Perceptions and expectations of autonomous vehicles – a snapshot of vulnerable road user opinion. *Technological Forecasting & Social Change*, *143*(October 2018), 9–13.

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in web questionnaires. *Developments in Applied Statistics*, *19*, 159–177.

Rinker, T. W. (2021). sentimentr: Calculate text polarity sentiment. New York: Buffalo. Version 2.9.0.

Rusch, M. L., Schall Jr, M. C., Gavin, P., Lee, J. D., Dawson, J. D., Vecera, S., & Rizzo, M. (2013). Directing driver attention with augmented reality cues. *Transportation Research. Part F, Traffic Psychology and Behaviour*, *16*, 127–137.

Schütze, H., & Silverstein, C. (1997). *Projections for efficient document clustering*. *ACM SIGIR forum: Vol. 31SI*. NY, USA: ACM New York (pp. 74–81).
Teodorovicz, T., Kun, A. L., Sadun, R., & Shaer, O. (2022). Multitasking while driving: A time use study of commuting knowledge workers to assess current and future uses. *International Journal of Human-Computer Studies*, *162*, Article 102789.
Teodorovicz, T., Sadun, R., Kun, A. L., & Shaer, O. (2021). *Working from home during COVID19: Evidence from time-use studies*. Harvard Business School.
Thomas, K. L., Peter, W. F., & Darrell, L. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*, 259–284.
Varghese, V., & Jana, A. (2018). Impact of ICT on multitasking during travel and the value of travel time savings: Empirical evidences from Mumbai, India. *Travel Behaviour and Society*, *12*(December 2017), 11–22.
Wiedemann, G., & Wiedemann (2016). *Text mining for qualitative data analysis in the social sciences: Vol. 1*. Springer.
Young, S. D., Adelstein, B. D., & Ellis, S. R. (2007). Demand characteristics in assessing motion sickness in a virtual environment: Or does taking a motion sickness questionnaire make you sick? *IEEE Transactions on Visualization and Computer Graphics*, 422–428.