Uncovering Perceived Identification Accuracy of In-Vehicle Biometric Sensing

Abdallah El Ali
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
abdallah.el.ali@cwi.nl

Robert Zwitser
Faculty of Social and Behavioural Sciences
University of Amsterdam
Amsterdam, The Netherlands
r.j.zwitser@uva.nl

Liam Ashby
Centrum Wiskunde & Informatica
Faculty of Social and Behavioural Sciences
University of Amsterdam
Amsterdam, The Netherlands
ashbyliam40@gmail.com

Pablo Cesar
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
p.s.cesar@cwi.nl

Andrew M. Webb
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
amwebb@acm.org

Abstract
Biometric techniques can help make vehicles safer to drive, authenticate users, and provide personalized in-car experiences. However, it is unclear to what extent users are willing to trade their personal biometric data for such benefits. In this early work, we conducted an open card sorting study (N=11) to better understand how well users perceive their physical, behavioral and physiological features can personally identify them. Findings showed that on average participants clustered features into six groups, and helped us revise ambiguous cards and better understand users’ clustering. These findings provide the basis for a follow up online closed card sorting study to more fully understand perceived identification accuracy of (in-vehicle) biometric sensing. By uncovering this at a larger scale, we can then further study the privacy and user experience trade-off in (automated) vehicles.

Author Keywords
Biometrics; sensing; card sorting; perceived accuracy; privacy; in-vehicle

CCS Concepts
+Human-centered computing → HCI theory, concepts and models;
Introduction
Biometric techniques can make vehicles safer to drive (e.g., drowsiness detection [44]), protect them against theft (e.g., tap-based authentication [20]), provide higher cost efficiency (e.g., rewarding good driver profiles through insurance telematics [11]), or provide personalized in-car experiences (e.g., route personalization based on driving style [6]). These advantages and benefits notwithstanding, biometric techniques and the promise of connected cars also raise large privacy concerns [6, 45], for example concerning the privacy and security of (sensed) personal data [35].

Prior research has addressed these issues by focusing on the so-called privacy-personalization paradox, namely that consumers who value information transparency are also less likely to participate in personalization [2]. This phenomenon has been generally less studied in an automotive context, and for automated driving in particular. The latter context becomes especially relevant, where non-driving-related activities (e.g., texting, eating) are desirable to perform in the car [34], which brings about opportunities for a wider range of in-vehicle biometric sensing [36]. To address the question of trading personal biometric data for in-vehicle user benefits, we take the first step here using an open card sorting study to better understand how well users perceive their physical, behavioral and physiological features can personally identify them. We contribute early empirical findings of how participants interpret and cluster such features.

Background and Related work
According to Jain et al. [15], biometric recognition “can be defined as the science of establishing the identity of an individual based on the physical and/or behavioral characteristics of the person either in a fully automated or a semi-automated manner.” Furthermore, the context in which data is collected, and who the data is to be sold to (e.g, research institute, commercial company), have been found to impact users’ willingness to disclose personal information [38]. Past research has also investigated how different types of personal information are valued differently by users. For example, amongst various types of data (i.e. application, location, communication), location data has been found to be the most valuable to users [40].

Within the automotive domain, there has been prior work on automotive activity recognition, for example using capacitive proximity sensing [4], that does not rely on collecting highly personal user data. Other work addresses the privacy implications of accelerometer data, where geographic tracking of drivers becomes possible using only a phone’s accelerometer sensor [18]. However, there is still no systematic analysis of how and when drivers and passengers willing to share their personal, biometric data, in exchange for a better in-vehicle trip experience.

Methods
Open card sorting
Open card sorting is a widely-used method (in web design) to create taxonomies based on users’ groupings of the content. Importantly, it can be effective for discovering the optimal organization of information according to users’ viewpoints [48]. Furthermore, it has been shown to have high cross-study reliability [16]. Since we are concerned with biometric technologies and perceived identification accuracy, card sorting is a suitable method to better understand how users cluster physical, behavioral and physiological features that such technologies can sense.

Card selection process
To understand how users perceive different biometric features, we first collect a common, relevant set of features,
Table 1: Cards (features), example biometric technology, and citation. For this study, we only tested cards 1-32 (plus "My signature").

<table>
<thead>
<tr>
<th>#</th>
<th>Card</th>
<th>Example Technology</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My face</td>
<td>Face recognition (2D,3D)</td>
<td>[31, 35]</td>
</tr>
<tr>
<td>2</td>
<td>My facial expressions</td>
<td>Facial emotion recognition</td>
<td>[31, 5]</td>
</tr>
<tr>
<td>3</td>
<td>My ears</td>
<td>Ear identification</td>
<td>[22]</td>
</tr>
<tr>
<td>4</td>
<td>My eyes</td>
<td>Eye tracking; iris and retina recognition</td>
<td>[31]</td>
</tr>
<tr>
<td>5</td>
<td>My physical activity</td>
<td>Physical activity recognition (accelerometers)</td>
<td>[24]</td>
</tr>
<tr>
<td>6</td>
<td>My fingerprints</td>
<td>Fingerprint recognition</td>
<td>[31, 35]</td>
</tr>
<tr>
<td>7</td>
<td>My walking style</td>
<td>Gait recognition</td>
<td>[31]</td>
</tr>
<tr>
<td>8</td>
<td>My hands</td>
<td>Hand geometry recognition (2D,3D); Palm and finger vein recognition</td>
<td>[31, 35, 37]</td>
</tr>
<tr>
<td>9</td>
<td>My sleeping patterns</td>
<td>Sleep classification</td>
<td>[21]</td>
</tr>
<tr>
<td>10</td>
<td>My smell</td>
<td>Odor recognition</td>
<td>[31]</td>
</tr>
<tr>
<td>11</td>
<td>My handwriting</td>
<td>Optical Character Recognition (OCR); signature recognition</td>
<td>[26, 31, 8]</td>
</tr>
<tr>
<td>12</td>
<td>My touches on a smartphone</td>
<td>Touchscreen dynamics</td>
<td>[31]</td>
</tr>
<tr>
<td>13</td>
<td>My mouse movements</td>
<td>Mouse movement dynamics</td>
<td>[31]</td>
</tr>
<tr>
<td>14</td>
<td>My typing on a keyboard</td>
<td>Keystroke recognition</td>
<td>[31]</td>
</tr>
<tr>
<td>15</td>
<td>My voice</td>
<td>Speaker identification, verification, authentication</td>
<td>[10, 35]</td>
</tr>
<tr>
<td>16</td>
<td>My teeth</td>
<td>Teeth recognition</td>
<td>[19]</td>
</tr>
<tr>
<td>17</td>
<td>My footprints</td>
<td>Footprint (size and shape) recognition</td>
<td>[28]</td>
</tr>
<tr>
<td>18</td>
<td>My heartbeat</td>
<td>Electrocardiogram - Heart Rate Variability</td>
<td>[31]</td>
</tr>
<tr>
<td>19</td>
<td>My writing style</td>
<td>Stylometry</td>
<td>[31]</td>
</tr>
<tr>
<td>20</td>
<td>My smartphone app usage</td>
<td>App usage fingerprints</td>
<td>[42]</td>
</tr>
<tr>
<td>21</td>
<td>My posture</td>
<td>Posture recognition</td>
<td>[33]</td>
</tr>
<tr>
<td>22</td>
<td>My media listening history</td>
<td>Personalized music emotion recognition</td>
<td>[49]</td>
</tr>
<tr>
<td>23</td>
<td>My media watching history</td>
<td>Profiling TV viewers using data mining</td>
<td>[39]</td>
</tr>
<tr>
<td>24</td>
<td>My driving style</td>
<td>Driver and driving style recognition</td>
<td>[43]</td>
</tr>
<tr>
<td>25</td>
<td>My interaction patterns with an in-car information system</td>
<td>Modelling driver IVIS interactions</td>
<td>[12]</td>
</tr>
<tr>
<td>26</td>
<td>My SMS messages</td>
<td>User classification based on SMS messages</td>
<td>[14]</td>
</tr>
<tr>
<td>27</td>
<td>My locations on a given day</td>
<td>Location tracking</td>
<td>[7]</td>
</tr>
<tr>
<td>28</td>
<td>My hand sweat</td>
<td>Galvanic Skin Response for emotion recognition</td>
<td>[23]</td>
</tr>
<tr>
<td>29</td>
<td>My hand gestures</td>
<td>Hand gesture recognition</td>
<td>[32]</td>
</tr>
<tr>
<td>30</td>
<td>My electrical brain activity</td>
<td>Electroencephalography: fNIRS</td>
<td>[41, 36]</td>
</tr>
<tr>
<td>31</td>
<td>My breathing</td>
<td>Breathing monitoring</td>
<td>[30]</td>
</tr>
<tr>
<td>32</td>
<td>My genetic makeup</td>
<td>DNA matching</td>
<td>[31]</td>
</tr>
<tr>
<td>33</td>
<td>My muscle movements</td>
<td>Electromyography motion classification</td>
<td>[29]</td>
</tr>
<tr>
<td>34</td>
<td>My personality</td>
<td>Personality classification</td>
<td>[9]</td>
</tr>
<tr>
<td>35</td>
<td>My eye gaze patterns</td>
<td>Eye tracking and user identification</td>
<td>[31, 13]</td>
</tr>
<tr>
<td>36</td>
<td>My driving route</td>
<td>Personalized route recommendation</td>
<td>[22]</td>
</tr>
<tr>
<td>37</td>
<td>My company in a vehicle</td>
<td>Person identification using social connections</td>
<td>[47, 46]</td>
</tr>
<tr>
<td>38</td>
<td>My body temperature</td>
<td>Thermal imaging</td>
<td>[1]</td>
</tr>
<tr>
<td>39</td>
<td>My eating style</td>
<td>Eating episode classification (wearable sensors)</td>
<td>[3]</td>
</tr>
<tr>
<td>40</td>
<td>My saliva</td>
<td>Stress (from cortisol) recognition</td>
<td>[17]</td>
</tr>
</tbody>
</table>

Procedure

Participants were tested in a lab environment. They were provided with an information sheet, filled and signed a consent and participant information form, then given a task demonstration. They were provided with sorting instructions on the desk (Figure 1). Instructions stated: “Please group the cards into distinct sets by how accurately someone can identify you using only the feature stated on a card.” We deliberately did not constrain participants to consider they were inside a vehicle for two reasons: (a) to gain a general understanding of perceived identification accuracy (b) to not burden participants to reflect on both automated and non-automated driving. After they sorted the cards, they were asked to label each group. There was no time limit, nor a limit to the number of groups they can create. Participants were encouraged to think aloud, and asked to explain their final groupings. Each session was audio recorded, and lasted approximately 10-30 min.

Figure 2: Set of sorted cards by a participant.

It frames in an understandable manner. This means deliberately not mentioning the underlying biometric technology, and instead only presenting the physical, behavioral or physiological feature that a technology processes. For example, to investigate facial recognition, we present participants with the card “My face”. Within an (automated) driving context, the space of biometrics is quite vast [36]. We did an extensive literature search by querying the ACM Digital Library and Google Scholar for papers published on biometric techniques, across all contexts (automotive, mobile, health, etc.). We then selected any biometrics that could in principle be also used in an (automated) vehicle. Our initial search resulted in 33 cards, which we test in this study. Our ongoing search resulted in eight more features (33-40)², which we aim to test in a follow up study. All cards are shown in Table 1. While not exhaustive, this list is sufficient for investigating perceptions of common biometrics.

²“My signature” was discarded, as explained later.
Participants
We recruited 11 participants (6f, 5m), aged between 21-38 (M=25.6, SD=5.4). Eight were students, and the remainder graduate-level or higher. Seven stated they had a technical background. Three participants had driving experience, two were learning how to drive, and the rest no experience. Participants did not receive monetary compensation.

Early Results and Discussion
Grouping, labeling, and task instructions
The mean group size created was 5.8 (SD=1.1). Given the question we asked, participants largely created a ranked list of groups, ranging from high (e.g., P2: "That’s me!"); P1: "Bio features") to low perceived identification accuracy (P9: "Things I do not or do almost the same as other people"; P5: "No one can identify me (not enough info)"). Overall, participants’ labels were consistent, however differed by grouping size. Given the range and mean group size, we aim to test six categories for our future study.

To better understand how our cards were grouped, we used Ward's hierarchical agglomerative clustering method [27] with k=6 clusters (creating six groups: G1-G6), where the resulting dendrogram using Jaccard similarity is shown in Figure 3. The dendrogram shows that some while groups contain physical features (e.g., G1: genetic makeup, fingerprints, eyes, ...), others are geared towards interactions with technology (e.g., G3: mouse movements, touches on smartphone, ...). The last grouping (G6) contains inconsistent topics, where some are quite identifiable characteristics of individuals (e.g., facial expressions), whereas others less so (e.g., driving style). This leads us to consider adding a "Not applicable" (N/A) category. Finally, one participant raised the question of whether a person or machine is doing the identification. Given this, we will adjust our future instructions to reflect a focus on machine (computer) sensing.

Ambiguous and redundant cards
Not all the cards were immediately understandable to participants (e.g., P2 was confused by physical activity and genetic makeup). Due to this, we need to provide additional examples (beyond what is listed in Table 1) to decrease ambiguity. P3 was unsure why we have cards for "My face" and "My facial expressions" – while this may create redundancy, we find it is better to maintain separation with respect to the underlying recognition technology. Moreover, cultural factors were surfaced, for example concerning writing style (P2: "In Chinese I have a very special writing style...when people see my homework, they know it is me"). We also merged "My signature" with "My handwriting", which was found to be redundant. Finally, some cards were deemed inapplicable, e.g., if participants did not have driving experience (P9: "I can’t drive...kind of just eliminated that") – this further necessitates an N/A category.

Next Steps
Our open card sort study findings helped shape a follow up closed card sorting study, where we have built our own online tool to collect closed card sort data (Figure 4). We have modified our instructions and the number of closed categories according to insights gathered. We also now include the complete 40 card list, and provide examples for most cards to avoid ambiguity. Based on the closed card sort data, we will choose the extreme ends of identifiability categories, and use those for a follow up study on identifying (automated) in-vehicle user experience privacy tradeoffs (e.g., trading heartbeat data for alertness monitoring).

---

Footnote:
\(^3\)Age statistics are based on 10 participants only, as one participant chose not to disclose their age.
REFERENCES


