Author(s): Silvennoinen, Johanna; Kujala, Tuomo; Jokinen, Jussi P.P.

Title: Semantic distance as a critical factor in icon design for in-car infotainment systems

Year: 2017

Version:

SEMANTIC DISTANCE AS A CRITICAL FACTOR IN ICON DESIGN FOR IN-CAR INFOTAINMENT SYSTEMS

Abstract

In-car infotainment systems require icons that enable fluent cognitive information processing and safe interaction while driving. An important issue is how to find an optimised set of icons for different functions in terms of semantic distance. In an optimised icon set, every icon needs to be semantically as close as possible to the function it visually represents and semantically as far as possible from the other functions represented concurrently. In three experiments (N = 21 each), semantic distances of 19 icons to four menu functions were studied with preference rankings, verbal protocols, and the primed product comparisons method. The results show that the primed product comparisons method can be efficiently utilised for finding an optimised set of icons for time-critical applications out of a larger set of icons. The findings indicate the benefits of the novel methodological perspective into the icon design for safety-critical contexts in general.

Research highlights

· Semantic distance was operationalised as preference ratings and reaction times.

· In optimised icon set, each icon is semantically as close as possible to its function.

· In optimised icon set, each icon is semantically as far as possible from other functions.

· Preferences and verbal protocols alone could not discern the optimised icon set.

· The method was effective in detecting the optimised icon set from a larger icon set.

1   INTRODUCTION

As vehicle technology evolves, the complexity and connectivity of in-car infotainment systems continually increase. This surge in technology means that the driver increasingly has access to a large number of novel in-car online applications, which can offer improved communication, entertainment, route finding, as well as other useful in-car services on the road. One unfortunate downside of this progress, however, is the increased potential for drivers to be distracted from the safety-critical primary task of driving while utilising the services (Victor, Dozza, Bärgman, Boda, Engström, & Markkula, 2014; Klauer, Dingus, Neale, Sudweeks, & Ramsey 2006).

The evolution of in-car systems has led to a large growth in system functions and, along with this, a growth in visual icons that represent these functions. Furthermore, as novel applications are introduced into vehicle systems, easily distinguishable new icons are needed to represent these functions. In the driving context, a two-second glance off road can already be risky (Liang, Lee, & Yekhshatyan, 2012), which means the driver should be able to find and locate the desired function from the in-car menus as fast as possible. This leads to novel challenges for the in-car interface designers to find an optimised combination of such menu icons that can be recognised with a brief in-car glance (Dobres, Chahine, & Reimer, 2017; Dobres, Reimer, Mehler, Chahine, & Gould, 2014). Thus, effective icon design that enables fluent communication in human-computer interaction (HCI) is especially critical for interactions with in-car infotainment systems while driving.

In this interaction context, time is of the essence owing to the pressure to return eyes on the road. An action to be conducted by selecting an icon can be demanding owing to the competition of attention by the other icons on the display. Therefore, the focus of this paper is to examine the cognitive processing fluency of icons' semantic distance, and the relationship between an icon's visual representation and its intended meaning. Previous research has mainly focused on studying the semantic distance of individual icons (e.g., Isherwood, 2009; McDougall, Curry, & Bruijn, 1999). However, icon menus always include sets of icons, whose meanings are required to be distinguishable from the meanings represented by the other icons in the same icon set. Every icon in a menu needs to be semantically as close as possible to its intended function while also being semantically as far as possible from other icons' functionalities in the same icon set, so that the driver can recognise and select the required function safely while driving. Here, our aim is to present and validate a methodology to investigate and optimise icons' semantic distances in safety-

critical user interfaces, and thus to provide insights into icon design for safe interactions while driving.

In order to find an optimised set of icons for time-critical applications out of a larger set of icons, we first explored four sets of possible icons and their semantic distances to four different in-car navigation system functionalities by studying participants' preference rankings and their verbal protocols. To examine how quickly these preferred icons can be processed, in the second experiment, we tested how quickly people are able to make the preferential judgments concerning the icon functions. Finally, in order to find set of icons where the icons of different functions are easily distinguishable, in the third experiment, we tested how quickly users identify icons of a given function when compared to icons of a different function. As a result, we present an icon set for the given functions, optimised for being individually quick to interpret, by referring to their intended meaning, as well as by being distinguishable as the icon of their intended function in the complete icon set.

2   ICON DESIGN IN THE AUTOMOTIVE CONTEXT

Icons stand for the objects they represent, that is, the displayed features and properties in icons resemble or imitate the objects they signify (Peirce, 1986). Icon metaphors are often elicited from real objects to emphasise familiarity (Blackwell, 2006), and in technological artefacts, can be defined as graphical representations that symbolize actions in technological environments (Ware, 2004). Icons are powerful elements in visual communication (Poulin, 2011) and enable users to accomplish technological tasks visually (Kay, 1990). Properly designed icons reduce system complexity and mental workload (Gittins, 1986), and provide better cognitive affordances than textual user interfaces (Garcia, Badre, & Stasko, 1994). Moreover, the large extent of icon-based user interfaces highlights visual icon design, not only to enhance communicability, but also to match user preferences (Huang, Shieh, & Chi, 2002).

Additionally, effectively designed iconic representations make objects, concepts, and actions easier to find, recognise, remember, and learn (Lidwell, Holden, & Butler, 2011). Thus, icons are more universally recognised than textual information (Lodding, 1983), are recognised quickly (Caplin, 2001), and are well remembered (Weidenbeck, 1999). Therefore, icons can be perceived immediately and enhance fluent communication and visual usability of interactive systems. This perceptual immediacy enables well designed icons to be grasped and understood effortlessly (Mullet & Sano, 1995) and the graphic representation of an icon affects its recognition rate and, therefore, influences user perception (Gatsou, Politis, &

Zevgolis, 2012). Immediate recognition and long memorability of icons raise challenges for efficient icon design. In practice, the intended functions of the icons might gain different meanings across users (Bocker, 1996; Isherwood, 2009; Isherwood, McDougall, & Curry, 2007), because icons convey semantic information through visual language that does not rely on strict rules in the same way as written words (Carr, 1986). Further, icons follow less strict rules than written language, which also contributes to their ambiguity between individuals.

Several studies have focused on visual icon characteristics and design principles in general (e.g., Byrne, 1993; Frutiger, 1997; Gaver, 1990; Gittins, 1986; Goonetilleke, Shih, On, & Fritsch, 2001; Ng & Chan, 2008). For example, some cognitive features in icon effectiveness include familiarity, concreteness, visual complexity, meaningfulness, and semantic distance (McDougall et al., 1999; Ng & Chan, 2008). Familiarity refers to the frequency of encounters with icons, concreteness to the abstraction level of the icons visual representation, complexity to the number of visual elements in the icon, and meaningfulness is how the icon's meaning is perceived (Ng & Chan, 2008). In addition, several icon design principles, aiming towards cognitive processing fluency, have been presented. For example, immediacy refers to effective recognition and cognitive processing fluency, in which the design focus is on the most essential visual elements through simplification and abstraction, not merely reducing the elements (Mullet & Sano, 1995). Icon design should follow the principle of generality by representing a broader category (e.g. painting supplies) of the idea, rather than an exact object (i.e. detailed photographic representation of some specific paint roller) in a cohesive manner within an icon set. Characterisation is utilised to emphasise the most essential features of a representation, including the most advantageous viewpoint. To design for communicability, knowledge of the users, culture, and context of use is required (Mullet & Sano, 1995). In addition to these icon design principles, understanding of cognitive processing fluency of icons' semantic distance is needed to design for safe interactions while driving. Cognitive effectiveness of semantic distance has not been studied in terms of icon sets, merely concerning individual icons, and thus, icon design principles would need to include this viewpoint of semantic distance, especially in time- and safety-critical interaction contexts.

For visual information processing to be fluent and effective, pictorial representations must activate correct mental models that match the representation's function (Isherwood, 2009). In icon design, this relationship is called semantic distance, a necessary factor in cognitive effectiveness of icon interpretation (Isherwood et al., 2007; Isherwood, 2009; McDougall et al., 1999; McDougall, Curry, & de Bruijn, 2001; McDougall & Reppa, 2013;

Ng and Chan, 2008). However, methodological approaches to semantic distance research have not addressed the role of semantic distance in a set of icons, or the requirements that a specific application context can set. Icon sets for specific interaction contexts have been studied in relation to, for instance, transportation and leisure activities (Prada, Rodrigues, Silva & Garrido, 2015), emergency medical information systems (Salman, Cheng, & Patterson, 2012), and user interfaces for pre-schoolers (Chiu, Koong, & Fan, 2012). A few studies have concentrated on icon design and testing in the automotive domain (e.g., Johann & Mahr, 2011). There are general guidelines for in-car user interface icons based on human factors principles and standards (e.g., ISO 15008 2009), but these are typically limited to enabling legibility and clarity of the icons while on the move. Thus, icon design research lacks studies of users' interpretations and semantic meanings of visual icon design in in-car infotainment systems for icon sets in which individual icons' semantic distances can be recognised quickly.

Recently, this interaction context has become a significant challenge for visual designers because of the explosion of in-car functionalities and services that are made available to the driver (e.g., Norman, 2007). This stresses the requirement that all the different functions available in the in-car infotainment system should have descriptive and intuitive icons communicating meanings unambiguously. Icons are required to be designed as enabling interactions with in-car systems as efficiently as possible in order to minimise the potential for distraction while driving (NHTSA, 2013). In this time- and safety-critical interaction context, milliseconds can truly make a difference. User interfaces for in-car infotainment systems in particular require icons for which the semantic distance to the associated functions are as close as possible. The driver should be able to locate and select the correct function within a brief in-car glance.

According to the analysis and the early visual sampling model of Wierwille (1993), drivers prefer to keep off-road glance durations on average between 0.5 and 1.6 s depending on the demands of the driving situation. In addition, naturalistic driving studies have found significant statistical associations between safety-critical incident risk and the off-road glance duration. According to Liang et al. (2012), the risks start to significantly increase with off-road glances that last more than 2 s. A subsequent analysis on the same 100-car study data by Liang, Lee, and Horrey (2014) suggested the general risk threshold is even lower, at 1.7 s, that is, near the 1.6-s upper limit of the Wierwille's (1993) model. Thus, semantic distance research in the automotive domain needs to take into account the cognitive processing fluency of icons in terms of reaction times in selection tasks as well as drivers' subjective

preferences. We suggest that results of effective processing fluency can be obtained by merging reaction times with preference rankings of subjective significances of the icons' functions. Preference construction is highly context sensitive and influenced by users' goals (Warren, McGraw, & van Boven, 2011). In this study, the factor influencing user preferences is the in-car navigation system's icon's semantic distance, that is, the relatedness of the visual representation and its intended function.

In this paper, we introduce and study a method intended to enable in-car user interface designers to find an optimised set of menu icons with optimal semantic distances from a large set of alternative icon designs. Recently, Dobres et al. (2017; 2014) introduced a similar method for finding an optimal typeface for in-car infotainment systems to provide the best legibility of digital text on in-car displays. The focus of the current study is to resolve an optimised visual icon design set for an in-car navigation system menu with primed product comparisons, based on user preferences and reaction times (Jokinen, Silvennoinen, Perälä, & Saariluoma, 2015). An optimised combination of icons for this specific design context requires optimal semantic distances. For an optimised icon set, the semantic distance needs to function effectively between one icon and its intended meaning, and also between different icons and their meanings, so that the icons differ from one another enough to optimise the selection of correct icon from the set of icons.

3   EXPERIMENT 1: RANKINGS AND VERBAL PROTOCOLS

The purpose of Experiment 1 was to explore four sets of possible icons and their semantic distances to four different in-car navigation system functions by studying participants' preference rankings per function and the associated verbal protocols. By studying the verbal protocols behind the preference rankings, we aim to indicate the significance of the context-specific semantic distance for icon design and to better understand its role when compared to the other icon design principles. In addition, the preference rankings act as a comparison point for further data gathered with the primed product comparison method.

3.1   Method

3.1.1   Participants and Stimuli

Participants ($N = 21$) were recruited for the experiment (11 male and 10 female) via student email lists of University of Jyväskylä. The primed product comparison method has been validated with 20 participants (Jokinen et al. 2015), and this sample size was used as a general guideline in the experiments. The mean age of the participants was 24.3 ($SD = 5.2$,

age range 20–40). All the participants had previous experience with navigation systems, and driving experience for at least either two years or 20 000 km.

The icons were selected from two sources for obtaining the stimuli for the experiment. In total, 18 icons were included into Experiment 1, as displayed in Table 1. Eight icons were acquired from a commercial in-car navigation system under development (HERE Auto, https://en.wikipedia.org/wiki/Here_(company)#Here_Auto). The X-icon was excluded from Experiment 1 owing to its inappropriate conventional meaning to represent any of the four functions. However, the X-icon was later utilised as a validity check for the primed product comparisons method in Experiment 2. An additional 11 icons were designed for this experiment for comparison purposes. The new icons were designed according to icon design principles of immediacy, generality, cohesiveness, characterisation, and communicability (Mullet & Sano, 1995).

(Table 1. around here) Table 1. *Icons used as stimuli in the experiments.*

Icons from the commercial navigation system	X-icon	Road signs	Mag. glass	Flag	Menu	Star	Gear	Jar		
New Icons	Pen enve-lope	Pen	Road signs	Bino-culars	White mag. glass	Wren-ch	Points menu	Screw driver	Star folder	Hea-rts	POI stars

Icon metaphor conventions in navigation system user interfaces and other software were also taken into account. Additionally, the icons were designed according to the style of the icons from the commercial navigation system. The style of the existing icons was mainly based on two-dimensionality, simplicity, consistency, and achromatic colour scheme, and it followed design conventions of pictograms. The new icons were designed to evaluate users' preferences and interpretations of conventions; preferred level of simplicity; and combinations of metaphor conventions. These were examined in terms of users' interpretations of icons' semantic distances in in-car navigation system user interfaces.

3.1.2   Procedure

The experiment started with participants ranking the icons using the given navigation system functions as criteria. Participants were asked to select one icon as the first option to the best

match to the given function, then to select a second-best option to express the function in question, then the third, fourth, fifth and sixth. The functions were 'Enter address', 'Search', 'Settings', and 'My destinations'. Both the commercial and newly designed icons were used (except the X-icon). Combinations of icons concerning different functions are presented in Table 2. 'Search' and 'Settings' included five icons each owing to conventional status of the selected icons to represent these two functions. 'Enter address' and 'My destinations' included six icons each in order to examine more options in terms of semantic distance owing to the lack of an established status of these explicit terms to represent the functions.

(Table 2. around here) Table 2. *Icon ranking for the four functions.*

Functions	Icons					
Enter address						
Search						
Settings						
My destinations						

The participants were shown one of the four icon sets at a time, and the function above the icons on a 22-inch 1650 x 1050 px display. The size of the icons was 57 x 72 px. Ranking was chosen as the method instead of scoring the icons for representing a function (on a scale), because ranking as a non-parametric method enables clearer results in the case of a small sample size. The participants were asked to think aloud while ranking the icons in order to extract verbal protocols (Ericsson & Simon, 1980; Boren & Ramey, 2000).

3.1.3 Data Analysis

Icon ranking data were analysed to detect which icons are the most preferred in relation to the semantic distance of the function and the icon. Ranking of the icons was conducted by labelling the best option with number 1, second best with number 2, and so forth. The total rank scores from the icon ranking task were used to compare the icons with each other. The Friedman test was used to test if the ranks were statistically significant from each other.

The thinking aloud data was transcribed into a textual format and analysed with qualitative content analysis (Krippendorf, 2004) utilising an *interpretation framework* that

defines the objects found in the data on the conceptual level, and through which the results of this experiment are produced (Silverman, 2005). The conceptual core of the interpretation framework was based on detecting semantic distances between the proximity of the relationship between the visual representation of an icon and the function it is intended to represent. In addition, icon design principles of familiarity, concreteness, visual complexity and meaningfulness served as concepts in the interpretation framework. The analysis consisted of familiarization, organization, and categorization of the data. The goal of the analysis was to understand the reasons behind user preferences and interpretations of the icons' meanings and functions.

3.2   Results

Ranking of the icons resulted in the following order for the four functions (Table 3). In all the icon rankings, the mean ranks were different from one another, as suggested by statistically significant Friedman tests, which were for 'Enter address' $\chi^2(5) = 20.6$, $p = .001$, 'Search' $\chi^2(4) = 67.9$, $p < .001$, 'Settings' $\chi^2(4) = 62.9$, $p < .001$, and for 'My destinations' $\chi^2(5) = 62.2$, $p < .001$. Lower mean rank indicates higher preference. The tables also include information of how often specific icons were selected as the first option.

(Table 3. around here) Table 3. *Ranking of the icons for the four functions.*

Enter address	Pen	Mag. glass	Road signs	Pen envelope	Flag	Arrow sign
Mean rank *(SD)*	2,1 (SD = 1.5)	3,1 (SD = 2.0)	3,7 (SD = 1.6)	4,5 (SD = 1.8)	3,4 (SD = 1.4)	4,1 (SD = 0.9)
Selected 1. x times	10	6	2	2	1	0

Search	Mag. glass	White mag. glass	Bino-culars	Road signs	Flag	
Mean rank *(SD)*	1,4 (SD = 1.0)	1,8 (SD = 0.4)	3 (SD = 0.4)	4,2 (SD = 0.7)	4,7 (SD = 0.5)	
Selected 1. x times	16	5	0	0	0	

Settings	Wheel	Wrench	Screw driver	Points menu	Menu	
Mean rank *(SD)*	1,3 (SD = 0.6)	1,8 (SD = 0.9)	3,5 (SD = 0.5)	4,1 (SD = 0.7)	4,3 (SD = 0.8)	
Selected 1. x times	16	5	0	0	0	

My desti-nations	POI stars	Star	Star folder	Flag	Hearts	Jar
Mean rank *(SD)*	2 (SD = 0.8)	2,3 (SD = 1.3)	2,6 (SD = 1.3)	4,5 (SD = 1.4)	3,9 (SD = 1.0)	5,7 (SD = 0.7)
Selected 1. x times	6	8	5	2	0	0

The pen icon was selected as the first option to represent 'Enter address'. The descriptions of selecting the pen included comments such as: *"pen symbolises entering something like writing something"* and *"because it´s for writing, I think. For me it´s the clearest, because you have to type the address and actually write it"*. The functionality was emphasised literally in resemblance to writing and the concreteness of the icon was emphasised. Entering was related to writing and writing to typing. Even though writing in navigation systems is not

10

done with an actual pen, the metaphor of a pen as a writing tool was preferred due to concrete juxtaposition of real world objects and functions. The icon with an envelope and a pen also represented writing or entering something. However, this icon was selected as the last option because of its strong conventional status as an icon for sending email. Thus, contextual familiarity influenced the ranking of the pen envelope icon.

The magnifying glass icon was consistently preferred to represent 'Search' due to its familiarity and conventional status. Preferences were described, such as: *"I think it´s so common in referring to search, search in internet or in navigation system, so it´s the best"*. No other icons were ranked as the first options. The magnifying glass (with black inside area of the glass, i.e. Mag. glass) was preferred the most owing to its simplicity, concreteness, clarity, and good contrast. In addition, the black inside area of the glass was seen to reflect that the search has not yet been done. The white inside area of the white magnifying glass, was seen to communicate that the search has already been done, which could be utilised in indicating the stage of a search process. The white magnifying glass was chosen as the second option owing to its lack of simplicity and concreteness: *"the one with the white background, the same story but it´s a little bit more detailed and it´s harder to see it fast I think"*. The binoculars icon was often selected as the third option because it also refers to looking and finding something with a meaningful semantic distance.

The wheel icon was the most preferred icon for 'Settings'. Sixteen participants chose the gear icon as the first option, owing to its convention as a settings icon, familiarity from other software, and metaphor of adjusting something. The gear, wrench, and screwdriver icons were seen to belong to the same tool category. However, the gear was selected as the best option, for instance, with the following words: *"I was struggling with these two (gear & wrench). It…allows you to manipulate the feeling of such a system, but it is more pleasantly expressed, because the wrench here implies that I´m an engineer and all the settings would be for engineers"*. The gear icon carried subtle nuances in representing 'Settings' which were not conveyed through the wrench and the screwdriver icons. The remaining two icons were seen to relate to menu icons and were therefore not suitable for 'Settings' in in-car user interfaces.

The point-of-interest (POI) star icon was selected as the first option to represent 'My destinations'. The intended function of the icon was to access visited and stored favourite destinations. The POI sign was familiar from digital maps, and the POI signs were seen to resemble balloons or tear drops upside down. The star represented the meaning of a favourite. Participants combined these two signs into one understandable and meaningful icon

metaphor. Preferences were described, for instance with the following words: *"it has the star in it, so it refers to my favourites and also the background, the icon is used similarly in navigation systems, where this icon would be set as a marker somewhere"*. It was also stated that the POI star icon was preferred because it has multiple POI elements, which represents that there are many destinations, not only one destination. The star icon was ranked as the second-best option but chosen as the first option more often than the POI star icon. Overall, participants preferred icons with stars over the icon with a folder and a star because those were considered to be too complex and cluttered. The last options were the flag and the hearts icons. The flag was seen more like a destination marker and the hearts icon was seen as an unfamiliar icon in comparison to the icons with stars. The jar icon was selected as the worst owing to its lack of comprehensibility in the in-car navigation system context. It was interpreted to represent, for example, an on-off switch, battery, trash bin, memory, kitten angel, and a seat belt, without clear relation to its functionality.

Overall, the participants expressed frustration if the icons were not easily recognisable, and if they could not arrive into a sensible interpretation of the icon's representation to its meaning within the first interpretation. The first ranking indicated that the participants' impression of the icon meaning functioned as a strong predictor of the intended function in the ranking the icons while thinking aloud. If participants were hesitant about the meaning, they were not willing to pursue interpreting the icons. Frustration in interpreting the icons was expressed, for example, when interpreting the jar icon, the following words were used: *"I have no other clear implication what the kitten angel icon resembles to me"*, and when interpreting the road signs, the following words were used: *"I haven't seen it, it could be...I don't know, do I really have to say?"*.

3.3   Discussion

The rankings show that the pen icon for 'Enter address', magnifying glass icon for 'Search', gear icon for 'Settings', and star icon for 'My destinations' were the most preferred icons to represent these four functions. These icon metaphor conventions from other information systems software were interpreted as meaningful and understandable in in-car navigation systems. In line with these findings, in-car user interface design guidelines (e.g., NHTSA, 2013) recommend the use of internationally agreed upon standards or recognised industry practice relating to icons and symbols. However, conventional design does not automatically contribute to effective design (McDougall & Curry, 2004). Thus, icons' semantic distances need to be investigated in novel interaction contexts to understand whether the semantic

distance elicits the required mental models for the intended actions in the specific context of use, and how quickly the icons can be recognised among other icons. For example, in the 'Enter address' icon rankings, the interpretations of the pen envelope icon indicated the influence of conventions, familiarity, and context. Therefore, the context in which icons are to be applied acts as a significant determinant and modifier in interpretations of semantic meanings.

Conventions function through familiarity, which are learned from corresponding products. Besides familiarity, products that include something new are preferred if the combination of familiarity and novelty is optimal. The key is in providing something new while preserving familiarity (Hekkert, 2006). The balance between novelty and familiarity was encountered in the users' preferences of the integrated POI star icon, which combined elements from two different visual design contexts in one icon. Users were able to interpret the conventional star and the cartographic POI mark together and process the new integrated icon with meaningful semantic distance. According to this result, in-car navigation systems could benefit from a specific set of icons that combines conventional metaphors from operating systems and, for instance, cartographic signs.

In HCI, confusing interaction design leads to frustration and stress (Rogers, Sharp, & Preece, 2011), which also affects icon processing fluency. If the semantic distance of an icon metaphor and its intended function is not understood, users become frustrated quickly and lose interest in trying to interpret the icon, which underlines the importance of understanding users' interpretations of icons, and what kinds of actions are mentally represented. Insights into icon design with subtle nuances can be gained with user studies on preferences and verbal protocols associated to these. For instance, this study informed the design metaphor to be used for 'Settings'. The gear icon was considered suitable to represent 'Settings', in that it represents universalistic design, without implicating specific levels of expertise. New integrated icons for a specific interaction context can enhance intuitive interaction between users and technology, but they need to be designed according to the icon design principles and tested with user studies.

4   EXPERIMENT 2: PREFERENCES AND REACTION TIMES

In the first experiment, the participants ranked the icons into a preferential order in relation to four different in-car navigation system functions. However, especially in safety-critical design contexts, preference is not the only criterion for good icon design. In addition, the user must be able to quickly make the intended interpretation, which leads to the required action.

Therefore, a second experiment was designed in order to test how quickly people are able to make preferential judgments concerning icon functionalities with primed product comparisons (Jokinen et al. 2015) and if it takes less time to make the judgment for the more preferred icons when compared to the less preferred icons. The basic idea of the method is that the participant is first shown a prime, such as a function that an icon intends to refer to, and then two stimuli, such as two icons, are shown from which the participant then needs to choose the one that they prefer more, given the prime (Figure 1). The participant is asked to make this preferential judgment as quickly as possible, and the task is repeated many times with different combination of primes and stimulus pairs. The resulting data contains prime-specific preferences as well as reaction times, indicating how quickly the participants were able to make the comparison.

(Figure1. around here)
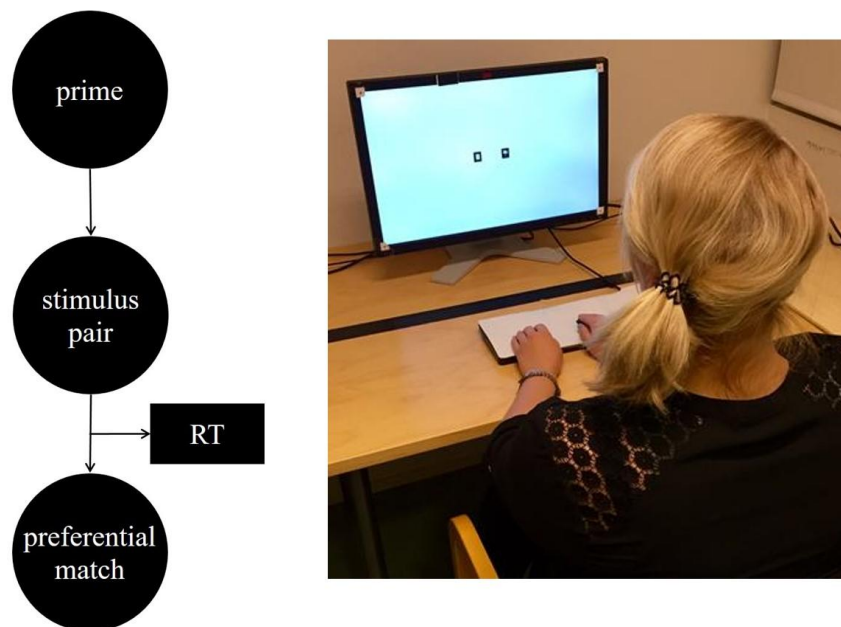
Figure 1. The procedure of primed product comparison method and experimental setup. In order to validate the icon preferences obtained in the first experiment in a more time-constrained context, we first hypothesise that:

H1. Preferences from the comparisons tasks correlate with icon rankings of Experiment 1.

Further, we propose that the comparison judgments should be conducted quickly. Because the method of primed product comparisons (Jokinen et al. 2015) cannot be directly used to analyse processing times of single stimuli, we use this experiment to explore the reaction times associated with pairwise icon comparisons. In the analysis, we focus on the upper threshold of 1600 ms by using Wierwille's (1993) visual sampling model, that is, an icon should be identifiable during a brief 1.6-s (maximum) in-car glance time. The reaction

14

times do not correspond directly to in-car glance times in the real world but we wanted to have a plausible maximum acceptable limit for a reaction time of a pairwise comparison. Our focus here was to find the optimised icon for each function in terms of semantic distance. There should be icons that are faster to process, and thus, we should see variance between reaction times for different icon pairs:

H2. There are differences in mean reaction times between icon comparisons.

In addition, we suggest that preference is at least partly dependent on the speed with which the participant is able to give a preferential match between a function and an icon, and thus:

H3. More preferred icons are selected faster than the less preferred icons.

4.1 Method

4.1.1 Participants and Stimuli

Participants ($N = 21$), 11 male and 10 female, were recruited with the same requirements as those in the Experiment 1: all had previous experience with navigation systems and driving experience of either at least two years or at least 20 000 km. Driving experience in terms of monthly kilometres was also asked before beginning the experiment. In addition, the participants were required not to have taken part in Experiment 1, because familiarity with the icons would have influenced the reaction time data. Participants' mean age was 28.5 (*SD* = 4.7, age range = 22–39).

The icons from Experiment 1 were reused as stimuli. The icons were presented on the same display as that used in Experiment 1 with custom software designed for the primed product comparison method. The participants' task in this experiment was to compare two icons at a time. The participant's viewing distance from the display varied approximately between 70 and 75 cm. The horizontal visual angle between the icons varied between 5.7° and 6.1°, that is, more than 5°, which places them outside of parafovea, where visual acuity is very poor (Rayner, 1998). In other words, the participants were able to only accurately observe one icon at a time. However, the distance between the icons was kept small (4.3 cm) in order to enable fast eye movements between the icons.

4.1.2 Procedure

The procedure followed the method for primed product comparisons developed by (Jokinen et al. 2015). The participant sits in front of a computer screen and a reaction time keyboard with two buttons (as shown in Figure 1). First, the participants are presented a prime, which

can be any word. After a fixed time, a pair of stimuli is shown side by side, and the participant's task is to choose the one that matches better to the prime, which was shown before the stimulus pair.

In the experiment reported here, the primes were the four in-car navigation functions as in the experiment 1: 'Enter address', 'Search', 'Settings', and 'My destinations'. Each prime was associated with all possible pairwise combinations of the icons, meaning that the participants were shown, in random order, one of the four functions coupled with any two of the icons that were intended to represent that function, until all possible combinations of function and a pair of icons had been displayed. In addition, for each four functions, one icon not intended to refer to that function was added to each functionality group from the icons of the first experiment. These extra icons (one function) served as a validity check for the method: the non-fitting icon was hypothesised to be preferred the least from the group of icons associated with certain functionality.

For 'Enter address', the icon was the jar icon, for 'Search' the X-icon, for 'Settings' the star icon, and for 'My destinations' the gear icon. Thus, a single task consisted of one of the four functions (displayed for three seconds) and a pair of icons, from which the participants had to choose the one they preferred as being more associated with the given function. There were 72 tasks in total.

4.1.3 Data Analysis

The method of primed product comparisons provides two kinds of data. First, the preferential matches, made by the participants by choosing which stimuli match with which primes, can be used to calculate preference scores (PSs). These scores have a range between 0 and 1, and indicate the preference level, or 'proportion preferred', compared to the other stimuli on a given prime. For example, a preference score of 0.9 would mean that a particular icon was chosen 90% of the time, when compared with the other icons used in the study for that function.

A comparison of the PSs reveals which icons are most preferred for given functions. Thus, PSs can be correlated with the rank scores obtained from Experiment 1 to provide the validity for the preference results of the first study. Another interpretation of H1 is that the preferences of the icons in the two studies have a large shared variance (R^2), indicating that the ranking task in Experiment 1 and the primed product comparison task in Experiment 2 result in similar icon evaluations.

In addition to the PSs, the method of primed product comparisons provides reaction time data associated with different choices. Here, the analysis focuses on the reaction time differences between the icon pairs (H2). Faster judgment times when comparing two icons related to a given prime indicate that the icon is encoded quickly, providing support for the use of the icon in time-critical contexts, such as in-car navigation systems. The proposition here is that people favour icons, which can be quickly associated with given functionalities, and thus, comparisons of icons with large difference in PSs should be faster than those with similar PSs (H3). The hypotheses were tested using generalised linear mixed modelling, as suggested by (Jokinen et al., 2015). The dependent variable was reaction time in seconds, and icon pair was the independent variable. The analysis was conducted four times, separately for each function. The distribution of reaction time was observed to be a gamma distribution, with reaction times over 5.0 s deviating from the theoretical gamma distribution and thus excluded from the analysis as outliers. The software utilised in data analyses were R 3.1.3, IBM SPSS Statistics 22.0, and MS Excel 2010.

4.2 Results

The PSs of the icons within the four in-car navigation system functions are displayed in Table 4. Shared variance between icon rank scores from Experiment 1, and icon PSs from Experiment 2 were for 'Enter address' $R^2 = .11$, 'Search' $R^2 = .98$, 'Settings' $R^2 = .99$, and 'My destinations' $R^2 = .66$. This means that the icons for Search and Settings were rated very similarly between the two experiments. Icons for My destinations were also rated similarly, but not as strongly as for these two. Finally, there was very little shared variance between the icon scores for Enter address between the experiments.

(Table 4. around here) Table 4. *Icon preference scores and the order of preference (rank) from Experiment 2 compared to the ranks of Experiment 1.*

Function	Icon	EXPERIMENT 1		EXPERIMENT 2	
		Mean rank	Grand rank	PS	PS rank
Enter address	Arrow sign	4.1	5	.68	2
	Flag	3.4	3	.52	4
	Mag. glass	3.1	2	.37	6
	Pen	**2.1**	**1**	**.73**	**1**
	Pen envelope	4.5	6	.45	5
	Road signs	3.7	4	.57	3
	Jar	-	-	.18	7
Search	Binoculars	3	3	.65	3
	Flag	4.7	5	.28	5
	Mag. glass	**1.4**	**1**	**.86**	**1.5**
	Road signs	4.2	4	.33	4
	White mag. glass	1.8	2	**.86**	**1.5**
	X-icon	-	-	.03	6
Settings	Menu	4.3	5	.34	5
	Points menu	4.1	4	.35	4
	Screwdriver	3.5	3	.48	3
	Gear	**1.3**	**1**	**.90**	**1**
	Wrench	1.8	2	.84	2
	Star	-	-	.09	6
My destinations	Flag	4.5	5	.31	5
	Hearts	3.9	4	.71	3
	Jar	5.7	6	.23	6
	POI stars	**2**	**1**	**.84**	**1**
	Star	2.3	2	.53	4
	Star folder	2.6	3	.77	2
	Gear	-	-	.10	7

The grand mean reaction time of all primes and stimuli pairs across all participants was 1.61 s ($SD = 1.17$, skewness = 5.15), but for the analysis, reaction times more than 5.0 s were removed, resulting in mean reaction time of 1.51 s ($SD = 0.74$, skewness = 1.43). The hypothesis that there are different reaction times between icon pairs (H2) was tested separately for each function. For 'Enter address', there were no statistically significant differences between the icon pairs, unlike the case for the other primes, as evidenced by statistically significant F-tests in the multilevel model. Of interest are the fastest and slowest

comparisons: for example, under 'Search', the participants used the least time for evaluations containing one of the two magnifying glass icons, unless both were present. This correlates with the overall preference of the magnifying glass for 'Search', and supports H3 that preference is at least partly dependent on the speed with which the participant is able to give a preferential match between a function and an icon.

The shared variance between PSs and reaction times were for 'Enter address' R^2 = .34, 'Search' R^2 = .55, 'Settings' R^2 = .51, and 'My destinations' R^2 = .50, indicating that generally, about half of the reaction times was explainable by how clearly the preferential match between two icons could be made. For example, when comparing the gear and menu icons to represent 'Settings', only two participants preferred the menu icon. The mean reaction time for this task was 1061 ms. Conversely, the participants were divided when comparing points menu and menu (38% chose the latter), and the mean reaction time was 1959 ms.

Further, for each function, an icon associated with another function in Experiment 1 was included to serve as a validity check for the method (see Table 4, jar icon for 'Enter address', X-icon for 'Search', star for 'Settings', and gear for 'My destinations'). These icons were hypothesised to be preferred the least from the group of icons associated with a certain function. The results of the validity check were as hypothesised; these icons were rated as the last option and preferred the least in comparison to the other icons within a given function.

4.3   Discussion

Experiment 2 resulted in preference scores that generally correlated highly with the preferences of Experiment 1 (H1 supported), although there were low levels of shared variance between the preference rankings of Experiment 1 and the preference scores of Experiment 2 for the 'Enter address' function. Further, reaction times indicated that preference was associated with faster judgment times, indicating that more preferred icons are also faster to process visually and mentally (semantic distance, H2 and H3 supported).

However, this experiment did not analyse how well the icons work as a whole set of menu icons, because only some of the icons were displayed with certain functions. This means that it may still be possible that, when compiling the total menu icon set for all the necessary in-car navigation system functions, there may be conflicts in the semantic distances between icons and different functions (e.g., Experiment 1: magnifying glass for 'Enter address' and 'Search'). In order to test this, Experiment 2 must be extended so that the icons are compared to each other under all functions. The optimised set of icons is a combination of

icons with each having the best preference score for its own intended function and the fastest reaction time when compared to any of the other icons under this function. In the current context, the reaction times for each of the selected icons should also be preferably under 1600 ms (Wierwille, 1993).

5 EXPERIMENT 3: OPTIMISED ICON SET

In Experiment 3, our aim was to find a best possible icon set by (1) minimising the semantic distance between the icons and the functions they represent, and simultaneously (2) maximizing the semantic distance between the icons and the other functions they do not represent.

5.1 Method

5.1.1 Participants and Stimuli

Participants ($N = 21$), 11 male and 10 female, were recruited with the same requirements as those in the previous experiments: all had experience with navigation systems, and driving experience of either at least two or more years or at least 20 000 km. Participants' mean age was 24.8 (SD = 4.6, age range = 20–37). Participants were required not to have taken part in Experiments 1 and 2. All the icons from Experiment 2 were included (in total 19 icons). The icons were presented in the same display as in that Experiments 1 and 2.

5.1.2 Procedure

The procedure followed the same method as that in Experiment 2, primed product comparisons (Jokinen et al. 2015), and the experimental setup was also same as that in Experiment 2. However, the icons were not segregated by their function; instead, all icons were compared to each other under all four functions. Thus, the total number of trials was 180 (the number of all possible pairs from ten icons, 45, multiplied by the number of functions, 4).

5.1.3 Data Analysis

The PSs were calculated as in the previous experiment, but this time, for each of the four functions, each icon got a PS. The goal of the analysis was to find the best possible icon set, based both on how preferred the icons were for their own most preferred function as well as how distinguishable they were from icons preferred for other functions. For each function, only those icons with PS > .70 were chosen, as per the cut-off suggested by (Jokinen et al. 2015). As often with statistical cut-offs, the chosen value is based on convenience rather than

rigorous analysis: less than the cut-off would include too many 'preferred' items, whereas more than the cut-off would only list a few top items. A cross-tabulation of pairwise reaction times for all chosen icons results in a dataset, which can be used to find the optimised icon set, based both on PS and pairwise reaction time.

This search results in a set of icons that contains icons with higher than desired minimum PS (here .70), and which have the largest overall semantic distance from the functions represented by the other icons in the set. Thus, the result is not necessarily the icon set with smallest average pairwise reaction times, but it is the icon set with smallest average pairwise reaction time for icons with PS > .70 for their chosen function. This is because the concept of semantic distance is not limited only to processing times, but also involves subjective preference.

5.2   Results

The PSs of the icons within the four in-car navigation system functions are displayed in Table 5, with PSs over .70 highlighted.

(Table 5. around here) Table 5. *Icon preference scores in the Experiment 3 for each four functions.*

ICON Preference scores	FUNCTION			
	Enter address	Search	Settings	My destinations
Arrow sign PS	.64	.57	.48	.58
Hearts PS	.39	.33	.26	**.78**
Mag. glass PS	.68	**.96**	.24	.33
POI stars PS	.43	.39	.25	**.86**
Star folder PS	.43	.35	.53	**.84**
White mag. glass PS	.64	**.92**	.31	.33
Pen PS	**.92**	.60	.57	.34
Road signs PS	.67	.51	.50	.70
Gear PS	.11	.19	**.96**	.14
Wrench PS	.08	.19	**.88**	.10

For 'Enter address', pen was the clearly preferred icon, and no other icons made the cut-off. For 'Search', both magnifying glass icons were preferred over the other icons. For 'Settings', the gear and wrench icons were considered as the most suitable ones, and for 'My destinations', the preferred icons were hearts, POI stars, and star folder. Based on only the PSs, a user interface designer could now freely pick from these possibilities any set of icons to represent the user interface functions. Before this, however, one should consider that while

all the icons are preferred, some may be easier to distinguish from the icons which were preferred for the other functions.

The cross-tabulation for pairwise reaction times of the icons with PS > .70 is shown in Table 6, which can be used to search for the best possible icon set, considering both preference and how easily distinguishable they are from the other icons.

(Table 6. around here) Table 6. *Pairwise reaction times in milliseconds for icons with preference scores over .70. Pairings of the same functions are suppressed.*

		Enter address	Search		Settings		My destinations		
		Pen	Mag. glass	White Mag. glass	Gear	Wrench	Hearts	POI stars	Star
Enter address	Pen	-	1675	1466	1164	1268	1599	1448	1235
Search	Mag. glass	1066	-	-	857	914	1073	1030	984
	White mag. glass	1315	-	-	917	1171	895	1070	1091
Set-tings	Gear	1129	1061	963	-	-	1007	1009	983
	Wrench	1206	1081	1300	-	-	1001	1031	1243
My desti-nations	Hearts	1478	1504	1544	1297	1365	-	-	-
	POI stars	1324	1517	1549	1202	1281	-	-	-
	Star	1528	1270	1554	1362	1487	-	-	-

For example, when considering the pen to hearts for Enter address, the participants took on average 1599 ms to indicate their preference. Comparing this to the average reaction time for pen and star folder, 1235ms, reveals that the latter comparison is easier to make. A designer should choose the star folder over hearts to represent 'My destinations', because while both are preferred icons for the function, they differ in how well the user can tell them apart from the pen icon, when searching the user interface for 'Enter address'. The possible number of combinations to consider increases even with the relatively small number of candidates with PS > .70 for their respective functions. The easiest way to use Table 6 for design is to choose any set of icons that has no pairwise reaction times over a certain threshold, such as 1600 ms (Wierwille 1993) (in our experiment, this would exclude Mag. glass).

However, it is also possible to search the combination with the smallest average pairwise reaction times. A search through all the combinations yielded the following set of

preferred and distinguishable icons (Table 7): pen for 'Enter address', white mag. glass for 'Search', gear for 'Settings', and star folder for 'My destinations'. The average reaction time for this set of icons is 1226 ms, and the largest pairwise comparison reaction time is between the star folder and white mag. glass when considering 'My destinations', i.e., 1554 ms.

(Table 7. around here) Table 7. *The optimised icon set.*

FUNCTION	Enter address	Search	Settings	My destinations
ICON	Pen	White mag. glass	Gear	Star folder

5.3   Discussion

The results of Experiment 3 indicate that the primed product comparisons method can be efficiently utilised for reducing the space of possible icons for system functions based on users' preferences and reaction times, but also for finding the best possible combination of icons out of alternative designs for a menu with different functions. The observed differences between the most preferred icons per function in Experiments 1 and 2 compared to the final icon set based on Experiment 3 illustrates that it is necessary to not only look at the subjective preferences (Experiment 1) and the associated reaction times per function (Experiment 2), but also to compare all the icons with all the functions of the menu under design in order to find the optimised set of icons.

In Experiment 1, the magnifying glass icon was ranked as the icon to represent 'Search'. The pairwise comparisons in Experiment 2 resulted with the same scores for the magnifying glass icon and the white magnifying glass icon. Finally, Experiment 3 indicated that the white magnifying glass is the more effective one in the combined set of icons. Additionally, in Experiments 1 and 2, the POI star icon was selected as the icon with the most efficient semantic distance to its intended function: 'My destinations'. However, in Experiment 3, the star folder icon for 'My destinations' function had the shortest semantic distance to its intended function and the largest to the other icons representing the three other functions, and was thus selected for the final set of icons. A possible next step could be to further lower the larger reaction times of the best icon set by applying small changes to these icons by studying different icon design characteristics, such as colour as a pop-out effect within an icon set.

6 GENERAL DISCUSSION

In this paper, we have introduced and validated a method based on primed product comparisons (Jokinen et al. 2015) in the context of in-car interface icon design in order to enable an in-car user interface designer to find an optimised set of menu icons with optimal semantic distances from a large set of alternative icon designs.

In Experiment 1, we started by exploring drivers' preference rankings of four sets of icons for four in-car navigation system functions and the role of semantic distance behind the rankings by studying their associated verbal protocols. In Experiment 2, the high levels of shared variance between the preference rankings of Experiment 1 and the preference scores of Experiment 2 indicated that the primed product comparison method can provide highly similar results when compared to mere preference ranking. There was a low level of shared variance between the preference rankings and the preference scores on the 'Enter address' function ($R^2 = .11$), but for the other functions, the levels were high ($R^2 > .66$). The discrepancy on the 'Enter address' function suggests that there was more variance in the preferences for this function than for the others, which is in line with the findings of the both experiments. This may be explained by the lack of an established convention to represent the function. However, Experiment 3 was finally able to discriminate the optimal icon also for this function when all the icons were compared pairwise against all the functions.

The primed product comparison method provides additional information compared to mere preference rankings. The reaction times of Experiment 2 indicated that preference was associated with faster judgment times, indicating that more preferred icons are also faster to process visually and mentally (i.e., semantic distance is significantly associated with preference, and therefore efficiently operationalised with preference scores and reaction times). The results of Experiment 3 clearly indicate that the primed product comparisons method can be efficiently utilised not only for reducing the space of possible icons for system functionalities based on users' preferences and reaction times (as in Experiment 2), but also for indicating the best possible combination of icons for a menu with different functions out of many possible combinations based on semantic distances.

For time- and safety-critical contexts, in particular, such as in-car infotainment systems, the optimised combination should not be based only on users' preferences but also on processing times for associating the intended function to an icon while competing for attention with all the other icons visible on the display. An optimised icon set is one in which each icon is semantically as close as possible to the function that it visually represents while,

at the same time, as far as possible to the other functionalities represented at the same time in the user interface. The primed product comparison method was able to indicate this kind of set of icons out of a large number of possibilities (in this case, 19 icon alternatives to represent four functions).

Finally, the results of this study indicate that the following icons were the most optimal combination in terms of semantic distance for in-car navigation system user interfaces out of the icons under study; the pen icon for 'Enter address, the white magnifying glass as the icon for 'Search', the gear as the icon for 'Settings', and the star folder icon for 'My destinations'. It should be noted that these icons with the optimal semantic distances would have not been found and selected based on the results of Experiment 1 and/or Experiment 2 alone, but Experiment 3 was required in order to find this optimised combined set. All of the selected icons had the highest preference scores for their intended function, and the participants were able to make the preference judgement between an icon and its intended function in less than 1600 ms when displayed with any of the competing icons. For the in-car context, this time limit can be critical as it has been found to be the maximum time drivers prefer to look off road with a single glance in any traffic situation (Wierwille 1993). According to the analysis of (Liang et al. 2014), in-car glances more than 1700 ms long have a significant statistical association with safety-critical incident risk in real traffic. Despite of the lack of direct comparability between reaction times and in-car glance times, the findings suggest that the primed product comparisons method can be highly valuable for icon design in general, but for time-critical contexts in particular by minimising the required time to identify a menu function among a set of menu icons. This decreases the total glance time required to search a display, and may also decrease individual glance durations in glance-like information sampling conditions (Dobres et al., 2017).

6.1   Limitations and Future Research

The proposed method applies best for optimising the first-time contact with user interface icons. After a sufficient experience with a system, the users will probably become much more efficient in recognising the icons and processing times will decrease, as familiarity with pictorial representations ease cognitive information processing (e.g., McDougall & Reppa, 2013). However, for time- and safety-critical contexts, such as in-car systems, the user interfaces should be optimised for as fast adoption as possible. It can take a while until a set of icons with ambiguous semantic distances, within an icon and between the icons in the set, is efficiently memorised, especially if the use of the system functionalities is infrequent.

Future research should assess the relationship between semantic distance and learnability of the individual icons as well as the relationship between semantic distance and the efficiency of visual search of an icon among a combination of icons.

The number of in-car functions offered on modern in-car touch screen displays will continue to increase, and the greater the number of functions, the more important will be the optimisation of the user interface to reduce visual demands (Dobres et al., 2017; 2014). However, further research should validate the assumed positive effects of the optimised menu icons on visual distraction compared to the less optimal icon set, for instance, in a driving simulator experiment with secondary visual-search tasks. The reaction times in pairwise comparisons do not directly predict in-car glance times in the real world, as there were, for instance, no gaze movements from the forward roadway to the display and back simulated in our experiments. In addition, there are often more than only two icons displayed at a time in the menus of in-car navigation systems. However, we wanted to have a plausible maximum acceptable limit for a reaction time of a pairwise comparison, which was adopted from the visual sampling model of Wierwille (1993). Further experiments with visual-search tasks are necessary in order to evaluate if an icon for a given function in a menu of icons can be found in less than 1600 ms of in-car glance time. Whether concurrent cognitive load affects the search times and, thus, modifies what is the optimised icon set while the cognitive load is high, should be further studied.

Drivers tend to split in-car glances after a certain level of uncertainty of the driving environment is reached, for instance, if finding a menu item takes more than 1600 ms (Wierwille, 1993). Thus, one could argue that the icon processing or interpretation time is not that critical in this context. However, there is evidence suggesting that the durations of all the encoding steps required to complete an in-car task should be minimised in order to minimise the possibility of visual distraction (Kujala & Salvucci, 2015). There are a number of ways to decrease the processing time of an icon besides minimising the semantic distance to the intended function. The results of Experiment 1 suggest that new integrated icons for infotainment systems could be further elaborated. The participants were able to combine icon convention and cartographic symbols together easily and establish meaning for the new integrated POI star icon.

Several studies (e.g., Kujala & Salvucci, 2015; Lasch & Kujala, 2012; Kujala & Saariluoma, 2011) have indicated that limiting the number of concurrently displayed in-car menu items to six (or less) decreases the probability of long glances at the display. However, it is not unusual to see more than six menu items displayed on in-car displays at a time in a

modern in-car infotainment systems. The proposed method can be well utilised to optimise larger sets of icons, although according to these studies, it would make sense to design in-car menu structures with the maximum of six functions displayed simultaneously per screen, and optimise each menu icon set for a screen with the primed product comparison method. This would also enable faster tests for each screen compared to testing a screen with larger number of functions.

Generalisability of the optimised icon set to other icon design contexts can be partly considered. Icons for search and settings functions can be efficiently interpreted also in other than time- and safety-critical user interfaces owing to their general nature as menu functions. However, for example, the POI star icon with a cartographic sign could elicit confusing interpretations if attached to user interfaces without a cartographic use context. In this study, the focus was on semantic distances of icon metaphors, and the icons used as stimuli were black and white pictograms, which might have affected the processing times, because colour information draws attention and enhances memory performance more effectively than black and white information (Farley & Grant, 1976). Moreover, not only does the processing time spent to interpret the semantic distance of an icon guide the icon design decisions, so does the design context, which sets demands for visual usability (e.g., concerning legibility). Further analysis of icons' visual characteristics could focus on detecting different design features' effects on processing times, such as saliency effects. The method can be applied to study various different designs, but the variables studied need to be controlled in order to measure the effect of the characteristic under investigation, for example colour, abstractness of pictorial metaphors, and design eras of the icons (e.g. Silvennoinen & Jokinen, 2016).

In this study, the studied icons were selected to be simplistic with little number of variables in the icons' pictorial representations to focus on examining semantic distances of the icons. The primed comparison method, as described, is intended to find the optimised set of menu icons in terms of semantic distance. There could be other icon design principles (e.g., the principles presented in Section 2) based on which an icon set can be optimised. However, we have argued that for the context of safety-critical systems, semantic distance is a critical icon design factor.

Owing to practical reasons and depending on the number of alternative icon designs, it can be useful to run the paired product comparisons in two stages, in a similar approach to ours here: first, reducing the overall number of icons by testing separate sets per function; and then, testing against all the functions with the smaller (combined) set. Even if this type of testing in two stages with twenty participants can require 40 hours or more, the benefits for

the final product can be large compared to a design/decision process that would solely rely on the intuition of the designer(s). Owing to the ambiguity of visual language, which does not rely on strict rules as does written language (Carr, 1986), the intended functions for the icons can be interpreted with altering meanings (Isherwood, 2009; Isherwood et al., 2007). However, images are recognised and processed faster than textual information (Lodding, 1983; Lidwell et al., 2011), and thus, adding textual information to icons can reduce ambiguities but increase processing times. The method of primed product comparisons is beneficial to interaction designers in optimising the whole icon set, not just individual icons. In designing and renewing icon sets, or introducing new icons to an existing icon set, the method can be used to detect the processing times and preferences of the whole icon set.

In addition, user interfaces are not globally preferred similarly among different cultures, and the design decisions also affect usability of the systems (Reinecke & Bernstein, 2011). Thus, icons in user interfaces might convey different semantic information in different cultures. Localization of icon's semantic distance could be tested with primed product comparisons to obtain the most effective semantic distance between the icon's representation and its intended function for the target culture. In addition, age (Zahabi, Machado, Pankok, Lau, Liao, Hummer et al., 2017; Ortiz, Castro, Alarcón, Soler, & Anera, 2013), previous experiences (Chi & Dewi, 2014), and familiarity with technological devices can influence the interpretations and thus the results. The three experiments reported here were conducted with university students. The participants were deliberately recruited to be a homogenous group of participants for control purposes since this was the first time of studying the method in terms of close and far semantic distances of icons. Therefore, future studies will include more heterogeneous participant groups, for example elderly people.

The basic tools for the primed product comparisons are easy to implement, and the technological requirements are simple. Parts of the testing and analysis steps could also be automated to improve efficiency. For instance, the last step of pairwise comparisons, required for detecting the optimal combined icon set, could be done by an algorithm searching for the combination with smallest average pairwise reaction times, and thus, lowering the manual workload of the process.

Future research of icon design for in-car infotainment systems will greatly benefit from studying semantic distance as the relatedness of the intended function of the icon's pictorial representation, especially as related to users' mental models of the action, elicited. Additionally, the method could be applied to examine semantic distances of information obtained also with other sensory modalities than the visual modality. For instance, auditory

and haptic information of in-car functionalities could be examined in terms of semantic distance. Moreover, processing fluency of combinations of sound and visual information could be studied with semantic distance, as audio and visual information are the most common output channels in current in-car user interfaces (Zhao, Brumby, Chignell, Salvucci, & Goyal, 2013), and are thus required to be comprehended quickly. Although we have focused on in-car navigation system icons here, the current findings, method, and design principles are likely to be applicable for the design of in-vehicle user interfaces in general, and, for instance, on roadway sign design (Zahabi et al., 2017). A cumulating database on icon features' effects on semantic distances and the associated processing times in the automotive context could be utilised in in-car user interface designer tools, such as Distract-R (Salvucci, 2009), that are intended for rapid testing of distraction effects of in-car user interface designs.

ACKNOWLEDGEMENTS

References

Blackwell, A. (2006). The Reification of metaphor as a design tool. *ACM Transactions on Computer Human Interaction, 13*(4), 490–530. DOI:http://dx.doi.org/10.1145/1188816.1188820

Bocker, M. (1993). A Multiple Index Approach for the Evaluation of Pictograms. *Proceedings of the International Symposium on Human Factors in Telecommunication*, pp. 73–84.

Boren, M.T., & Ramey, J. (2000). Thinking Aloud: Reconciling Theory and Practice. *IEEE Transactions on Professional Communication, 43*(3), 261–278. DOI:http://dx.doi.org/10.1109/47.867942

Byrne, M.D. (1993). Using Icons to find Documents: Simplicity is Critical. *Proceedings of Conference on Human Factors in Computing Systems,* INTERACT '93 and CHI '93. ACM Press, New York, NY, 446–465. DOI:http://dx.doi.org/10.1145/169059.169369

Carr, T.H. (1986). Perceiving Visual Language. In: K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of Perception and Human Performance: Vol. 2. Cognitive Processes and Performance* (pp. 1–92). Oxford, UK: John Wiley and Sons.

Caplin, S. (2001). *Icon Design: Graphic Icons in Computer Interface Design*. London, UK: Cassell.

Chi, C. F., & Dewi, R. S. (2014). Matching performance of vehicle icons in graphical and textual formats. *Applied ergonomics, 45*(4), 904-916.

Chiu, S., Koong, C-S., & Fan, S-H. (2012). Icon Design for Preschoolers: Implications Derived From Child Development. *Procedia – Social and Behavioral Sciences, 64*, 228–237. DOI:http://dx.doi.org/10.1016/j.sbspro.2012.11.027

Dobres, J., Chahine, N., & Reimer, B. (2017). Effects of ambient illumination, contrast polarity, and letter size on text legibility under glance-like reading. *Applied Ergonomics*, *60*, 68-73. DOI:http://dx.doi.org/10.1016/j.apergo.2016.11.001

Dobres J., Reimer, B., Mehler B., Chahine, N., & Gould, D. (2014). A pilot study measuring the relative legibility of five simplified Chinese typefaces using psychophysical methods. *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,* AutomotiveUI '14. ACM press, New York, NY. DOI:http://dx.doi.org/10.1145/2667317.2667339

Ericsson, K.A., & Simon, H.A. (1980). Verbal Reports as Data. *Psychological Review, 87*(3), 215–251.

Farley F.H., & Grant, A.P. (1976). Arousal and cognition: Memory for color versus black and white multimedia presentation. *The Journal of Psychology: Interdisciplinary and Applied, 94*(1),147–150. DOI:http://dx.doi.org/10.1080/00223980.1976.9921410

Frutiger, A. (1997). *Signs and Symbols. Their Design and Meaning,* second ed. Cham, Switzerland: Syndor Press GmbH.

García, M., Badre A.N., & Stasko, J.T. (1994). Development and Validation of Icons Varying in Their Abstractness. *Interacting with Computers, 6*(2), 191–211. DOI:http://dx.doi.org/10.1016/0953-5438(94)90024-8

Gatsou, C., Politis, A., & Zevgolis, D. (2012). The importance of mobile interface icons on user interaction. *International Journal of Computing Science and Applications, 9*(3), 92–107.

Gaver, W.W. (1991). Technology Affordances. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology,* CHI'91. ACM press, New York, NY, 79–84. DOI:http://dx.doi.org/10.1145/108844.108856

Gittins, D. (1986). Icon-based Human-computer Interaction. *International Journal of Man-Machine Studies, 24*, 519–543. DOI:http://dx.doi.org/10.1016/S0020-7373(86)80007-4

Goonetilleke, R.S., Shih, H.M., On, H.K., & Fritsch, J. (2001). The effects of training and representational characteristics in icon design. *International Journal of Human-Computer Studies, 55*, 741–760. DOI:http://dx.doi.org/10.1006/ijhc.2001.0501

Hekkert, P. (2006). Design Aesthetics: Principles of Pleasure in Design. *Psychology Science 48*(2), 157–172.

Huang, S-H., Shieh, K-K., & Chi, C-F. (2002). Factors affecting the design of computer icons. *International Journal of Industrial Ergonomics, 29*, 211–218. DOI:http://dx.doi.org/10.1016/S0169-8141(01)00064-6

International Organization for Standardization. (2009). Road vehicles — Ergonomic aspects of transport information and control systems — Specifications and test procedures for in-vehicle visual presentation (ISO 15008:2009), ISO, UK.

Isherwood, S.J. (2009). Graphics and Semantics: The Relationship between What is Seen and What Is Meant in Icon Design. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics, Lecture Notes in Computer Science, Vol. 5639* (pp. 197–205). Berlin Heidelberg: Springer-Verlag.

Isherwood, S.J., McDougall, S.J.P., & Curry, M.B. (2007). Icon Identification in Context: The Changing Role of Icon Characteristics with User Experience. *Human Factors, 49*(3), 465–476. DOI:http://dx.doi.org/10.1518/001872007X200102

Johann V., & Mahr, A. (2011). New Icon Selection Test for the Automotive Domain Based on Relevant ISO Usability Criteria. *Adjunct Proceedings of 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2011.* Salzburg, Austria, 25–26.

Jokinen, J.P.P., Silvennoinen, J.M., Perälä, P.M.H., & Saariluoma, P. (2015). Quick affective judgments: Validation of a method for primed product comparisons. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '15. ACM Press, New York, NY, 2221–2230. DOI:http://dx.doi.org/10.1145/2702123.2702422

Kay, A. (1990). User interface: A personal view. In B. Laurel, & J.S. Mountford (Eds.), *The Art of human-computer interface design* (pp. 191–207). Boston, MA: Addison-Wesley Longman Publishing.

Klauer, S.G., Dingus, T.A. Neale, V.L. Sudweeks, J.D., & Ramsey, D.J. (2006). *The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data.* DOT HS Rep. 810 594. U.S. National Highway Traffic Safety Administration, Washington, DC.

Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*, second ed. Thousand Oaks, CA: Sage.

Kujala T., & Salvucci, D.D. (2015). Modeling visual sampling on in-car displays: The challenge of predicting safety-critical lapses of control. *International Journal of Human-Computer Studies, 79*, 66–78. DOI:http://dx.doi.org/10.1016/j.ijhcs.2015.02.009

Lasch, A., & Kujala, T. (2012). Designing browsing for in-car music player: effects of touch screen scrolling techniques, items per page and screen orientation on driver distraction. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 41-48). ACM.

Kujala, T., & Saariluoma, P. (2011). Effects of menu structure and touch screen scrolling style on the variability of glance durations during in-vehicle visual search tasks. *Ergonomics*, *54*(8), 716-732.

Liang, Y., Lee, J.D., & Yekhshatyan, L. (2012). How dangerous is looking away from the road? Algorithms predict crash risk from glance patterns in naturalistic driving.

Human Factors, 54(6), 1104–1116.
DOI:http://dx.doi.org/10.1177/0018720812446965

Liang, Y., Lee, J.D., Horrey, & W.J. (2014). A looming crisis: the distribution of off-road glance duration in moments leading up to crashes/near-crashes in naturalistic driving. *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting*. Sage, 2102–2106. DOI:http://dx.doi.org/10.1177/1541931214581442

Lidwell, W., Holden, K., & Butler, J. (2003). *Universal Principles of Design*. Beverly, MA: Rockport Publishers.

Lodding, K. (1983). Iconic interfacing. *IEEE Computer Graphics and Applications, 3*(2), 11–20. DOI:http://dx.doi.org/10.1109/MCG.1983.262982

McDougall, S.J.P., & Curry, M.B. (2004). More than just a picture: icon interpretation in context. *Proceedings of the International Workshop on Coping with Complexity*. University of Bath, UK, 73–81.

McDougall, S.J.P., Curry, M.B., & de Bruijn, O. (1999). Measuring symbol and icon characteristics: norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols. *Behavior Research Methods, Instruments, & Computers, 31*(3), 487–519.

McDougall, S.J.P., Curry, M.B., & de Bruijn, O. (2001). The effects of visual information on users' mental models: An evaluation of pathfinder analysis as a measure of icon usability. *International Journal of Cognitive Ergonomics, 5*(1), 59–84. DOI:http://dx.doi.org/10.1207/S15327566IJCE0501_4

McDougall, S.J.P., & Reppa, I. (2013). Ease of icon processing can predict icon appeal. In M. Kurosu (Ed.), *Human-Computer Interaction, Human-Centred Design Approaches, Methods, Tools, and Environments, Lecture Notes in Computer Science, Vol. 8004* (pp. 575–584). Berlin Heidelberg: Springer-Verlag.

Mullet, K., & Sano, D. (1995). *Designing Visual User Interfaces: Communication Oriented Techniques*. NJ: Prentice hall.

National Highway Traffic Safety Administration. (2013). Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. (NHTSA-2010-0053).

Ng, A.W.Y., & Chan, A.H.S. (2008). Visual and Cognitive Features on Icon Effectiveness. *Proceedings of the international multiconference of engineers and computer scientists,* IMECS 2008. Honk Kong, 19–21.

Norman, D. (2007). Three challenges for design. *Interactions, 14*(1), 46–47. DOI:http://dx.doi.org/10.1145/1189976.1190002

Ortiz, C., Castro, J. J., Alarcón, A., Soler, M., & Anera, R. G. (2013). Quantifying age-related differences in visual-discrimination capacity: drivers with and without visual impairment. *Applied ergonomics*, *44*(4), 523-531.

Peirce, C.S. (1986). Logic as semiotic: the theory of signs. In R.E. Innis (Ed.), *Semiotics: an introductory reader*. Bloomington, IN: Indiana University Press.

Poulin, R. (2011). *The language of graphic design: an illustrated handbook for understanding fundamental design principles.* Beverly, MA: Rockport Publishers.

Prada, M., Rodrigues, D., Silva, R.R., & Garrido, M.V. (2015). Lisbon symbol database (LSD): subjective norms for 600 symbols. *Behavior research methods*, 1-13. DOI:http://dx.doi.org/10.3758/s13428-015-0643-7

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. DOI:http://dx.doi.org/10.1037/0033-2909.124.3.372

Reinecke, K., & Bernstein, A. (2011). Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. *ACM Transactions on Computer-Human Interaction, 18*(2), Article 8, 29 pages. DOI:http://dx.doi.org/10.1145/1970378.1970382

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction Design: Beyond Human-Computer Interaction,* third ed. Chichester, UK: Wiley.

Salman, Y.B., Cheng, H.-I., & Patterson, P.E. (2012). Icon and user interface design for emergency medical information systems: A case study. *International Journal of Medical Informatics, 8*(1), 29–35. DOI:http://dx.doi.org/10.1016/j.ijmedinf.2011.08.005

Salvucci, D.D. (2009). Rapid prototyping and evaluation of in-vehicle interfaces. *ACM Transactions on Computer-Human Interaction, 16*(2), Article 9, 33 pages. DOI:http://dx.doi.org/10.1145/1534903.1534906

Silvennoinen, J., & Jokinen, J.P.P. (2016). Aesthetic Appeal and Visual Usability in Four Icon Design Eras. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM Press, New York, NY, 4390–4400. DOI:http://dx.doi.org/10.1145/2858036.2858462

Silverman, D. (2005). *Doing Qualitative Research*. Thousand Oaks, CA: Sage.

Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., & Markkula, G. (2014). *Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and*

Crash Risk. Technical Report No.S2-S08A-RW-1. Chalmers University of Technology, Sweden.

Ware, C. (2004). *Information Visualization*. CA: Morgan Kaufman.

Warren, C., McGraw, P.A., & van Boven, L. (2011). Values and preferences: defining preference construction. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(2), 193–205. DOI:http://dx.doi.org/10.1002/wcs.98

Weidenbeck, S. (1999). Use of icons and labels in an end user application program: an empirical study of learning and retention. *Behaviour & Information Technology, 18*(2), 68–82. DOI:http://dx.doi.org/10.1080/014492999119129

Wierwille, W.W. (1993). An initial model of visual sampling of in-car displays and controls. In A.G. Gale, I.D. Brown, C.M. Haslegrave, H.W. Kruysse, & S.P. Taylor (Eds.), *Vision in vehicles IV* (pp. 271–279). Amsterdam: Elsevier.

Zahabi, M., Machado, P., Pankok, C., Lau, M. Y., Liao, Y. F., Hummer, J., Rasdorf, W., & Kaber, D. B. (2017). The role of driver age in performance and attention allocation effects of roadway sign count, format and familiarity. *Applied Ergonomics*, *63*, 17-30.Zhao, S., Brumby, D. P., Chignell, M., Salvucci, D., & Goyal, S. (2013). Shared input multimodal mobile interfaces: Interaction modality effects on menu selection in single-task and dual-task environments. *Interacting with Computers, 25*(5), 386–403. DOI:http://dx.doi.org/10.1093/iwc/iws021