Further Evidence for Benefits of Verbal Route Guidance Instructions over Symbolic Spatial Guidance Instructions

Willem B. Verwey
TNO Institute for Human Factors
P.O. Box 23
NL 3769 ZG Soesterberg
The Netherlands
Phone: +31-3463-56439 FAX: +31-3463-53977
E-mail verwey@izf.tno.nl

ABSTRACT

This paper reports a study that investigated how to present simple guidance instructions in a car. The literature suggests an advantage for oral over visual information because the visual channel is already heavily loaded while driving. Earlier research suggested that this was due to the fact that verbal instructions are better suited than spatial symbols (arrows) on a visual display. The major aim of the present study was to replicate this finding and to test explanations for this phenomenon in order to come to more sophisticated guidelines. The results confirm the advantage of verbal instructions and suggest that it is caused by recoding of spatial guidance instructions into a verbal memory format which is not required for verbal instructions. It is concluded that navigation instructions should basically be oral although arrows or map-like images may be used to support oral presentation.

INTRODUCTION

In the near future, technological advances will allow a number of features to be included in the automotive environment that increase driver comfort and aid in vehicle operation. One of these features is electronic navigation support. The present paper addresses the issue how discrete route guidance instructions should be presented to minimize driver workload. Several researchers proposed to use the auditory modality for presenting guidance instructions rather than the visual one (e.g. [1][2]) because the visual modality is heavily loaded by driving (e.g. [3][4]) and should not be burdened by additional tasks. Verwey and Janssen [5] performed a field study in which similar route guidance instructions were presented in the visual and auditory modality and found that symbolic visual instructions (arrows) yielded almost twice as many navigation errors as oral instructions as well as a somewhat less smooth deceleration profile. This suggested that oral instructions do not only prevent overload of the visual channel but that these also are better suited for displaying this type of information. Since in the field study by Verwey and Janssen [5] instruction modality (auditory vs. visual) and format (verbal vs. spatial) were confounded, a laboratory study was performed which investigated whether the advantage for oral guidance instructions was due to the auditory modality or the verbal format of the guidance instructions [6]. The results confirmed the earlier obtained advantage for oral over visual spatial instructions and indicated that this advantage was not caused by the modality of presentation but, instead, by the verbal format.

The present study tested explanations for the advantage of oral guidance instructions in order to come to more advanced guidelines. A first explanation is that the use of only three arrows (two perpendicular, one straight) in the earlier studies had led to a recoding of the spatial instruction into a verbal format because the arrows did not indicate the angle of the side road and spatial aspects of lane keeping interfered with processing spatial information [7]. Then, the advantage of verbal information should disappear when the angles of the arrows always match the angle of the road indicated, that is, when they are congruent. A second explanation for the advantage of verbal instructions is that in the earlier studies spatial instructions were recoded into a verbal memory code because verbal codes in memory are less sensitive to decay over a longer interval than spatial codes [8]. Hence, spatial instruction may yield better performance when short intervals elapse between presentation of the instruction and confrontation with the intersection or cross roads.

These notions were tested in a laboratory experiment in which route guidance instructions were presented in either a verbal or a spatial format. All instructions were presented visually to prevent confounding of format and modality. After an interval of varied duration a schematic junction with four connecting roads was presented. It was the subjects' task to indicate whether the guidance instruction pertained to one of the connecting roads or not. Half of the subjects served in a pure condition in which the angle implied by the spatial and verbal guidance instruction matched the angle of the target road. The other subjects served in a mixed condition in which only simple instructions were used: 'left', 'right', and 'straight ahead'. So, in the mixed condition the angle implied by the guidance instruction was sometimes congruent with the angle of the side road and sometimes it was not.
Table I  The seven combinations of spatial and verbal route guidance instructions and target roads used in the pure condition. Each crossing also included two additional connecting roads.

<table>
<thead>
<tr>
<th>target road</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>spatial guidance instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>verbal guidance instruction</td>
<td>left</td>
<td>left</td>
<td>veering left</td>
<td>straight ahead</td>
<td>veering right</td>
<td>right</td>
<td>sharp right</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

METHOD

Route guidance instructions and road scenes

The pure condition contained only congruent instruction/road scene combinations. There were seven route guidance instructions, one for each multiple of 45° except 180°. In the spatial version of this condition each instruction consisted of an arrow under one of seven angles (Table I). In the verbal version the verbal equivalents of these arrows were presented.

In the mixed condition only three guidance instructions were used for each of the seven possible road directions. This yielded the nine combinations shown in Table II. Since a road under 45° to the right could be indicated by the instruction 'straight ahead' as well as by 'turn right' care was taken in the choice of the other side roads that each instruction unambiguously indicated only one road in the scene displayed.

Each traffic scene contained a grey colored road, which depicted the road on which the crossing was entered, and three connecting roads. The road scene was presented in birds-eye-view. An example is depicted in Fig. 1.

Table II The nine combinations of spatial and verbal route guidance instructions and target roads used in the mixed condition. The numbers 2, 5, and 8 are congruent instruction/road scene pairs, the others are incongruent.

<table>
<thead>
<tr>
<th>target road</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>spatial guidance instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>verbal guidance instruction</td>
<td>left</td>
<td>left</td>
<td>straight ahead</td>
<td>straight ahead</td>
<td>straight ahead</td>
<td>right</td>
<td>right</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>
Fig. 1 Example of the guidance instruction and road scene as presented in the spatial blocks of the pure condition. The hatched part represents the current road of the driver.

In positive trials the target road was selected by the computer and two additional roads were selected from the set of roads with an angle of more than 45° with the target road. This was done to prevent ambiguity about which road was indicated. In negative trials three roads with varying angles were randomly chosen from the set of roads that deviated more than 45° from the road that was indicated by the instruction.

Task

In all conditions, a spatial or verbal guidance instruction was presented for 1000 ms in the upper left hand corner of a monitor display. Immediately after the instruction had been removed from the screen a fixation dot appeared in the center of the screen for either 100 or 3000 ms. Next, the road scene was presented. The subjects' task was to indicate whether the guidance instruction indicated a road in the traffic scene that was displayed (pushing a rocker switch up) or whether it did not (pushing it down). In the example in Fig. 1 the road indicated by the instruction is present and the switch is pushed up.

Apparatus

An S-R interface with external clocks (accuracy 1 ms) connected to an IBM AT-3 with video digitizer (Matrox Inc.) controlled the timing of events, generated guidance instructions and recorded reaction times. An error was indicated by a 1000 Hz, 100 ms tone generated by a multi-purpose sound generator. The visual stimuli appeared on a 35 × 23 cm monitor (Barco, CDCT 2/51). The subjects sat alone in a sound-attenuated, dimly-lit 2 × 2 × 2 m cubicle (Amplisent) in front of a table with a small panel containing one rocker switch which could be pushed up or down. The stimuli appeared on the monitor about 110 cm from the subject’s eyes and could be perceived easily.

Subjects

Subjects were 8 male and 8 female university students between 18 and 27 years old who served as paid volunteers. All had normal or corrected vision.

Procedure

Upon arrival half of the subjects was randomly assigned to the pure condition, the other half to the mixed condition. Subjects were familiarized with the aim of the study and their task. In the pure condition the instruction was to always give a positive response when the guidance instruction exactly matched one of the side roads in the subsequently presented scene. In the mixed condition the angle of the side roads was not always equal to the direction implied by the guidance instruction and positive responses were required for the instruction/road scene pairs depicted in Table II.

Subjects performed 8 blocks of trials which lasted about 8 min when the interval between guidance instruction and road scene presentation was 100 ms and about 15 min when the interval was 3000 ms. There were 126 trials in each block, 63 positive trials and 63 negative trials. These trials were presented in random order. After 63 trials there was a short break of 20 s. In congruent blocks each of the seven positive instruction/road scene combinations was presented nine times whereas in mixed blocks each of the nine positive pairs was presented seven times.

Blocked independent variables were congruency (pure vs. mixed), interval duration (100 vs. 3000 ms), and instruction format (verbal vs. spatial). Congruency was varied between subjects, interval duration and instruction format/road scene within subjects. Pairs of guidance instructions and road
scenes formed another independent variable and was varied within blocks.

RESULTS

In all conditions, only the positive responses were analyzed. Initial analyses showed no difference between mirror images of instruction/road scene pairs and the data obtained in left and right versions of the same instruction/road pairs were pooled. Next, reaction times and arcsine transformed errors were analyzed in subject x instruction/road scene x interval duration x instruction format ANOVAs.

Initial analysis of the error data revealed a mean error rate of less than one percent and error data were not analyzed any further. Fig. 2 depicts the mean reaction times for the four instruction/road scene pairs in the pure condition.

Main effects were found of instruction/road scene \([F(3,21)=26.9, \ p<0.001]\), interval duration \([F(1,7)=17.1, \ p<0.01]\), and instruction format \([F(1,7)=6.38, \ p<0.05]\). The first main effect indicated that the reaction time increased when instructions and target roads had a bigger angle. A three-way interaction between instruction/road scene, interval duration, and instruction format \([F(3,21)=5.05, \ p<0.01]\) indicated that the main effect of interval duration was actually a result of the effect of interval duration in the verbal condition only. In Fig. 3 reaction times in the mixed condition are shown. The ANOVA showed main effects of instruction/road scene \([F(4,28)=25.7, \ p<0.001]\) and instruction format \([F(1,7)=18.5, \ p<0.01]\). No other effects were significant. A between-subject analysis on equal instruction/road pairs in the pure and mixed condition revealed no difference between pure versus mixed condition \((F<1)\) suggesting that the effects were not of a strategical nature.

![Fig. 2](image-url)  
**Fig. 2** Mean reaction times for the four instruction/road scene pairs in the pure condition.

![Fig. 3](image-url)  
**Fig. 3** Mean reaction times for the five instruction/road scene pairs in the mixed condition.
The present study replicated the advantage of verbal over spatial guidance instructions. The only exception to this were the results obtained with verbal instructions following the short intervals in the pure condition. There, responses to the instruction 'straight ahead' were very fast whereas responses to the instructions 'veering right' and 'veering left' were slow. Considering the length of each of the verbal instructions and taking the corresponding reaction times into account a plausible explanation for this pattern of results is that the 100 ms interval duration, after 1 s presentation of the instruction, was too short to fully process the verbal instructions and prepare for presentation of the road scene. So, this effect can probably be considered as an artefact of the procedure.

The advantage of verbal route guidance instructions was replicated in the other conditions despite the fact that there was no concurrent task with spatial features which could have played a role [7]. Also, recoding of spatial instructions was found to be a general phenomenon in that the disadvantage for spatial instructions was equally large in the pure and in the mixed conditions. This indicates that recoding of spatial guidance instructions takes place irrespective of the probability on incongruent instructions. Hence, the phenomenon is not a strategical effect.

The advantage of verbal instructions did not depend on interval duration so the data do not offer support for the notion that recoding is used to acquire a more stable memory code. Given the similar advantage of verbal instructions over spatial instructions of about 70 ms in all conditions, and the finding of equal effects of incongruent verbal and spatial instructions in the mixed conditions (Fig. 3), it appears that spatial instructions were translated to a verbal code only after the road scene had been presented. So, the data suggest that subjects retained the guidance instructions in the format of presentation. The finding that effects of congruency were equal in the spatial/mixed and verbal/mixed condition provides additional evidence for the notion that spatial instructions were recoded first and then treated similarly as verbal instructions.

In conclusion, the present study offers support for earlier findings that verbal route guidance instructions are more appropriate than symbolic, spatial instructions [5][6][7]. Since no other tasks was performed the recoding of spatial instructions can not have been used for preventing interference with spatial elements in the driving task or to come to more stable memory codes. Instead, recoding of arrows into a verbal memory format appears a robust phenomenon which is not caused by concurrent tasks and independent of the delay before application of the guidance instruction. Hence, by now there is converging evidence from various studies that route guidance systems should present discrete oral guidance instructions as the main source of information because (a) oral messages do not require the driver to scan a display and (b) oral messages are more appropriate for route guidance instructions than spatial instructions. Still, there may be two reasons to support oral guidance instructions by a simple visual indication. First, oral information is system-paced so the driver may miss information in critical situations and visual information can be used for confirmation. Then, visual and oral information should always include the same information to prevent the driver from looking for additional information at the visual display which is not given in the oral message. Second, simple verbal instructions may not be adequate for unambiguous presentation of guidance information in complex situations such as roundabouts. In that case, the driver may benefit from visually presented information, such as arrows or schematic map displays, because verbal guidance instructions become too complicated to comprehend quickly.

ACKNOWLEDGMENT

This study was supported by the Institute for Road Safety Research SWOV, Leidschendam, The Netherlands.

REFERENCES