Evaluation of legibility of not properly reflecting signs

de Waard, Dick ; Brookhuis, Karel; Mesken, J

Published in:
International Journal of Industrial Ergonomics

DOI:
10.1016/j.ergon.2005.01.008

Publication date:
2005

Evaluation of legibility of not properly reflecting signs

Dick de Waard, Karel Brookhuis, and Jolieke Mesken

Correspondence to
Dick de Waard, PhD,
Department of Psychology
University of Groningen
Grote Kruisstraat 2/1
9712 TS Groningen
The Netherlands
d.de.waard@rug.nl

Evaluation of legibility of not properly reflecting signs

Dick de Waard, Karel Brookhuis, and Jolieke Mesken

Abstract

In conditions of dew or temperatures below 0° C the reflection of reflecting material on signs can be suboptimal. Problems with respect to legibility under such conditions were studied, as well as subjective ratings of the signed information.
Images of the signs were presented using an LCD projector. Independent variables in the experiment were age, amount of information, stimulus quality, and presentation time. Dependent variables were errors in reproduction of the names of the cities, towns or villages on the signs, and subjective ratings. It turned out that all factors both separately and additively affected the amount of errors. Positive subjective ratings coincided mainly with level of reflection, and to a lesser degree with the amount of information, presentation time and familiarity of information.

Relevance to Industry

Different types of reflecting material have been developed to reflect light from signs optimally. Disadvantage of some of the materials is the effect dew and temperatures below 0 degrees centigrade can have on reflection and legibility. The study has practical consequences in terms of novelty and amount of information to be presented on such a sign.

Keywords
sign, legibility, elderly, simulation, dew, stimulus degradation

Introduction

Traffic signs are supposed to guide people to their destination. In unfamiliar environments signs are, up to date, the main source of information to determine where a driver is, and what direction he or she should follow. Even with the use of in-car navigation systems, on-site signs will still be used to verify route choice. Information on signs should be legible and clear at all times, i.e. during day and night. To take care that signs are legible during darkness while driving at high speed on a motorway, many signs have a layer of retroreflective sheeting so the sign reflects light emitted by vehicles' headlights. There are several factors that affect legibility of a sign. Amongst these are the font of typeset, the amount of information on the sign, the novelty of information (e.g., language), whether the sign is clean (signal-noise contrast, quality of the stimulus), and time required or available to glance at the information (cf. Purduski & Rys, 1999). Not only the sign properties, but also the processor of this information, the driver, is important. The state he or she is in, his or her visual acuity, attention (dual task performance), all affect ability to read the information displayed.

The present study was carried out as it was noted that there are signs treated with a layer of highly retroreflective material that sometimes have strongly reduced reflective properties. In general these are conditions of dew and a temperature below 0 degrees centigrade. Signs would either be fully affected leading to strongly reduced reflection, or could be partly affected giving the impression of being spotted and dirty. The question was whether these conditions would lead to problems with legibility and potentially missing of information. Main factor in the study was thus legibility quality (stimulus degradation), which has been shown to lead to increased reaction time and increased errors (Steyvers & Gaillard, 1993). Other factors that affect performance with respect to reading from a sign are familiarity of information, time available to perceive information, amount of information, and age of the driver. Older drivers' (mean age 65) dual tasks performance under time pressure has been shown to be worse than young drivers' performance (Brouwer et al., 1991), and with increasing age visual functions decline (Welford, 1985). It can thus be expected that elderly drivers will make more errors in reading the degraded signs than young drivers.

Method

Participants

A stratified sample (see table 1) aiming to have a total of at least 80 participants (20 per group) were recruited by an advertisement in a local newspaper, and from a pool of subjects that had participated in previous experiments. As experience may also play a role, a group of non-licensed young people was included.

>>> Table 1 about here

Stimulus material and design

As original material video recordings of traffic signs were taken that had been recorded at night. From this material three conditions were selected, firstly the fully reflecting sign, secondly a partly affected sign ("spotted", looking like it needs to be cleaned), and thirdly fully affected, strongly reduced in reflecting properties. These samples from the original video tape were electronic graphic files of signs were made that were projected by an LCD projector on a screen. The signs were the same in terms of colour (dark blue background, white letters) and font and font size as those that can be found along Dutch motorways. This font is developed by specifically four route guidance signs (see figure 1). Size of the projected signs was such that font size was optimal legible (font size in millimetres = 2.175 x watching distance (m) + 2.3, deduced from Sanders & McCormick, 1987). Factors that were varied on the signs were:

- *Sign Quality.* Three quality variations were used: clear, spotted (partly not reflecting, affected by dew), and dark (completely affected, not reflecting light properly)
- *Number of cities displayed.* Either one or three destinations were displayed on the sign.

- *Type of sign.* The sign either showed exit destinations, or destinations further along the motorway. Within the exit-signs *Familiarity* was a factor. Either existing cities or non-existing names of places were used. The latter were either variations on existing names (e.g., Venmond is a non-existing city and a mix of the existing cities of Venray and Helmond), or were variations on family names (e.g., Steyverdam is a variation on the last name Steyvers).
- *Presentation Time.* Was either brief (1 s) as glance time can expected to be in dense traffic, or normal (3 s).

Above factors resulted in 3 (Quality) x 2 (Number) x 3 (Type) x 2 (Time) = 36 different types of sign-presentation combinations. Factors were assigned to the stimulus material such that properties were evenly spread over the signs. Stimuli were put in random order and presentation order was a between-subject variable; either participants saw stimuli in that order or in reversed order. A total of 72 signs were presented. Figure 1 shows some examples of the stimulus material.

>>> Figure 1 about here

A new slide was announced by an alerting beep. Then the sign was projected for 1 or 3 seconds, and then participants were asked to turn a page of their answering form and to answer the question. If there were three names on the sign the target city was randomly assigned in advance (top, middle or bottom). Half of the questions were multiple choice (recognition), half were open end (recall). Participants had to cross out or write down the target city. After that for each sign four ratings on five-point Likert scales had to be given; a rating about legibility, pleasantness, clearness, and a rating in terms of good-bad. These scales were based on earlier experience with ratings of acceptance (Van der Laan et al., 1997)

Procedure

Six to twelve people were tested per experimental session. Before the experiment started it was assured that participants could read the font size by writing down projected car number-plates of font sizes increasing in seven steps. The mid stimulus was of the same font size as the traffic signs. After this, nine signs were presented and evaluated for practice purposes. Frequency of motorway use was assessed, and general opinion about "the blue guide signs" was asked.

Analyses

Answers were scored 'correct' or 'incorrect'. A new variable "error percentage' was calculated per factor (type of sign, presentation time, etc). Subjective scores were averaged for the three sign quality conditions. As statistical test a repeated measures analysis of variance was performed with SPSS/PC.

Results

A total of 105 participants participated in the experiment. Female participants were slightly in the majority, 58%. Other general information with respect to the participants can be found in table 2. Results with respect to motorway use and general opinion about the signs are summarised in table 3. On the basis of the results of the eye-sight test (errors in writing down two consecutive font size licence plate numbers equal or larger than the font size used on the signs) nine persons were excluded from analyses.

>>>> Tables 2 & 3 about here

Results from the questionnaire study confirmed that young drivers had driven least on motorways, both during day and at night. The difference in night driving between the high-experience middle aged driver group and the older drivers was only marginal. Relatively most complaints with respect to the signs came from the experienced drivers, one out of three reported that legibility was (sometimes) bad. The older drivers had relatively least complaints. Of all participants, five mentioned themselves as cause of these legibility problems, 14 participants blamed the signs. Other causes mentioned were: unclear font (6 x), dirty or improperly lit signs (both 4 x), fog or sunlight (3 x). Thirty four participants answered the question “under which conditions do you have difficulty reading the signs”. Eighteen mentioned problems in conditions of dazzling sunlight, nine during darkness, six during conditions of dew, and one did not give a specification. So 14% of the participants mentioned experience with reduced legibility in atmospheric conditions that could be associated with the reflective layer on the signs.

Errors- sign quality

Figure 2 shows the percentage of wrong answers given. Two effects can be seen, firstly errors differ per group ($F(3,92) = 8.96$, $p < 0.001$), and increase with age. Secondly, an increase of errors coincided with a decrease in reflecting properties of the sign (Hotellings $T^2 = 0.38$, $p < 0.001$). Best performance was found at a clear sign, worst performance at the dark sign. The interaction Group x Sign Quality was also significant (Hotellings $T^2 = 0.322$, $p < 0.001$).

>>> Figure 2 here

Errors - No. of cities displayed.

Not surprisingly, if the number of stimuli increases more errors are made. As the participants were not certain which of the three cities they were to reproduce or recognise, more errors were made with the three-city signs than with one-city signs ($F(1,92)= 174.6$, $p < 0.001$)). This is illustrated in figure 3. There was also a significant Group effect ($F(3,92)=8.96$, $p < 0.001$) and a significant Group x Number of cities interaction ($F(3,92) = 4.85$, $p < 0.01$). Older drivers made more errors, particularly if there were three cities.

>>> Figure 3 here

Errors - Type of sign

Three types were studied; ‘destination’-signs indicating destinations further along the motorway, familiar ‘exit’-signs (i.e. with existing cities and villages), and unfamiliar exit signs, with non-existing cities and villages. There were 24 signs of all three types. Figure 4 illustrates that least errors were made with destination signs, and most with unfamiliar exit signs (Hotellings $T^2 = 3.15$, $p < 0.001$). The motorway destination cities were relatively the largest cities and may therefore be best known. Univariate tests revealed significant differences on the dimensions ‘Destination vs. Exit’ ($F(1,92) = 9.97$, $p < 0.001$) and ‘Familiarity’ ($F(1,92) = 262.6$, $p < 0.001$). There was also a main effect of Group ($F(3,92) = 8.96$), and a significant Group x Sign type interaction (Hotellings $T^2 = 0.30$, $p < 0.001$). Univariate tests only showed significant results on Group x Familiarity, $F(3,92) = 8.94$, $p < 0.001$.

>>>> Figure 4 here

Errors - Presentation Time

A significant effect of Presentation time ($F(1,92) = 58.45$, $p < 0.001$) and of Group was found ($F(3,92) = 8.96$, $p < 0.001$). However, the interaction Presentation time x Group was not significant ($F(3,92) = 1.06$, NS), an indication that the older drivers did not perform additionally worse when exposed to the sign for a brief period.

>>>Figure 5 here

Errors – test method

Recognition versus recall was assessed by test method: multiple choice alternatives or open ended questions. In figure 6 the results are graphically displayed. More errors were made in the open-end condition ($F(1,92) = 64.5$, $p < 0.001$). Groups also differed significantly in error rate ($F(3,92) = 8.96$, $p < 0.001$), but the interaction Group x Test Method was not significant ($F(3,92) = 1.5$, NS).

>>> Figure 6 here

Subjective ratings – Sign quality

Signs were rated on 5-point Likert scales on the dimensions 'legible-illegible', 'pleasant-unpleasant', 'clear-unclear' and 'good-bad'. Scores were recoded to –2 .. +2, positive scores indicating a positive evaluation, and averages were calculated. Judgement is most negative for the spotted sign, in particular in terms of pleasantness. No effects of group were found on subjective ratings.

>> Table 4 about here

Discussion and conclusions

Traffic signs with reduced retroreflective properties, simulating the real world effects of signs affected by dew, lead to more errors in reproduction of the names indicated on these signs. Elderly drivers, aged 50 plus, make relatively more errors in name reproduction and recognition than younger drivers. The interaction (age)group by sign quality shows that this group in particular is vulnerable for the missing of information from affected signs. The start of this decrease in performance is already found in the middle-age group of experienced drivers (aged 25-45).
Further factors that were found to affect sign reading performance were amount of information on the sign, novelty of information, and stimulus presentation time (glance time). The effect of looking for familiar information, such as looking for the name of a place to exit the motorway; was studied both by test method (recognition vs. recall) and familiarity (existing name vs. non-existing name). Both factors were found to significantly affect performance: recognising or reading a familiar text is easier.

It is remarkable that no group effect of subjective ratings of the three sign qualities was found. As expected, elderly drivers perform overall worse in reading from affected signs, but they do not evaluate affected signs more negatively. All rate the clear sign as best, then the dark sign, then the spotted sign. Performance in terms of errors, however, is better for the spotted than for the dark signs!

Although the oldest group of drivers tends to avoid driving at night and in unfamiliar areas (e.g., Brouwer, 1996) the present study did not reveal differences in night-time driving for older drivers. In an ageing society, the industrial designer should therefore be particularly aware of elderly users' characteristics and attitudes (Graafmans et al., 1996) when developing products such as a retroreflective layer for traffic signs.

Problems with not-properly reflecting signs are likely to be related to specific situations. A limited time to read and process information may actually in practice be restricted as traffic density is mostly low at the time atmospheric conditions like dew affect sign reflection. However, if new crucial information is to be processed from a sign, low sign reflection may have a serious negative effect on legibility, as was shown in the present study.

Acknowledgement

This study was commissioned by the Dutch Ministry of Transport (AVV).

References

Brouwer, W.H., Waterink, W., Van Wolffelaar, P.C. & Rothengatter, J.A., 1991. Divided attention in experienced young and older drivers: lane tracking and visual analysis in a dynamic driving simulator. Human Factors, 33: 573-582.

Brouwer, W.H. (1996). Older drivers and attentional demands: consequences for human factors research. In: K. Brookhuis, C. Weikert, J. Moraal, and D. de Waard (Eds.). *Aging and Human Factors;* Proceedings of the Europe Chapter of the Human Factors and Ergonomics Society Annual Meeting 1993 in Soesterberg. Groningen, The Netherlands: Traffic Research Centre, University of Groningen. pp. 93 – 106

Graafmans, J.A.M., Fozard, J.L., Rietsema, J., Van Berlo, A., & Bouman, H., 1996. Gerontechnology: matching the technological environment to the needs and capacities of the elderly. In: K. Brookhuis,

C. Weikert, J. Moraal, and D. de Waard (Eds.). Aging and Human Factors; Proceedings of the Europe Chapter of the Human Factors and Ergonomics Society Annual Meeting 1993 in Soesterberg. Groningen, The Netherlands: Traffic Research Centre, University of Groningen. pp. 19 – 30

Purduski, J.M., & Rys, M.J., 1999. Evaluation of a new advance flagger traffic sign. International Journal of Industrial Ergonomics, 24, 107-114.

Sanders, M.S. and McCormick, E.J., 1987. Human Factors in Engineering and Design. New York: McGraw-Hill.

Steyvers, F.J.J.M., & Gaillard, A.W.K., 1993. The effects of sleep deprivation and incentives on human performance. Psychological Research, 55, 64-70.

Van der Laan, J.D., Heino, A., & De Waard, D., 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation Research - Part C: Emerging Technologies, 5, 1-10.

Welford, A.T. (1985). Changes of performance with age: an overview. In N. Charness (Ed.) Aging and Human Performance. Chichester: John Wiley & Sons.

Evaluation of legibility of not properly reflecting signs

Dick de Waard, Karel Brookhuis, and Jolieke Mesken

Table captions

Table 1. Participants recruited
Table 2 Group averages, Standard Deviation (sd) in brackets
Table 3. Frequency of motorway use in general and at night (%). Licensed drivers only. Lower part: opinion about signs (all participants)

Table 4. Average scores on legibility of signs (–2 is negative, e.g., illegible, +2 is positive, e.g., legible). In brackets the standard deviation is given

Figure Captions
Figure 1. Examples of the stimulus material.

Figure 2. Percentage errors made per sign quality and group (non = non-drivers (18-20 yrs), novice drivers (18-20 yrs), exper. = experienced drivers (25-45 yrs), and older drivers (aged 50+))

Figure 3. Percentage errors per group and number of cities displayed

Figure 4. Percentage errors made by information on sign and group. Fam=familiar, Unfam = unfamiliar.

Figure 5. Percentage errors made by sign presentation time (one or three seconds) and group

Figure 6. Percentage errors made by test method (open = open ended, MC = multiple choice) and group

Table 1. Participants recruited

Group	Age	Driving licence	Driving Experience (total mileage)
Non-drivers	18 – 20	No	None
Novice	18 – 20	Yes	Limited
Experienced	25 – 45	Yes	> 100 000 km
Older	50 +	Yes	Ample

Table 2 Group averages, Standard Deviation (sd) in brackets

	ALL	non-drivers	novice drivers	exper. drivers	older drivers
Number of subjects	105	25	28	24	28
Driving experience (years)	15.0 (14.9)	–	1.4 (0.7)	9.5 (5.0)	33.3 (8.4)
Annual mileage (km/year)	12800 (16400)	–	3200 (5500)	23600 (22000)	12800 (16400)
Eyes not corrected	48%	50%	54%	50%	39%
Wearing glasses	26%	13%	21%	13%	54%
Wearing contact lenses	26%	38%	25%	38%	7%

Table 3. Frequency of motorway use in general and at night (%). Licensed drivers only. Lower part: opinion about signs (all participants)

	In general				***At night***			
Frequency	ALL	Novice	Exper.	Older	ALL	novice	Exper.	Older
Daily	8 %	0 %	21 %	4 %	3 %	0 %	8 %	0 %
> 1 x week	28 %	4 %	42 %	41 %	22 %	4 %	38 %	26 %
> 1 x month	37 %	50 %	21 %	37 %	37 %	29 %	33 %	48 %
Max 1 x per month	28 %	46 %	17 %	19 %	37 %	64 %	21 %	22 %
(almost) never	0 %	0 %	0 %	0 %	3 %	4 %	0 %	4 %

General sign Legibility	ALL	Non drivers	Novice	Experienced	Older
Good	80 %	80 %	82 %	67 %	85 %
(sometimes) bad	20 %	20 %	18 %	33 %	15 %

Table 4. Average scores on legibility of signs (–2 is negative, e.g., illegible, +2 is positive, e.g., legible). In brackets the standard deviation is given

	Legible	Pleasant	Clear	Good
Sign Quality				
Clear	1.62 (0.36)	1.30 (0.48)	1.47 (0.41)	1.41 (0.49)
Spotty	0.63 (0.88)	–0.29 (1.01)	–0.10 (0.99)	–0.19 (1.05)
Dark	0.81 (0.84)	0.65 (0.93)	0.16 (0.95)	0.11 (0.98)

Sign quality: Clear **Sign quality: Spotted** **Sign quality: Dark**

Route (familiar only)

Exit – Familiar (existing)

Exit –Unfamiliar (non-existing)

Figure 1.

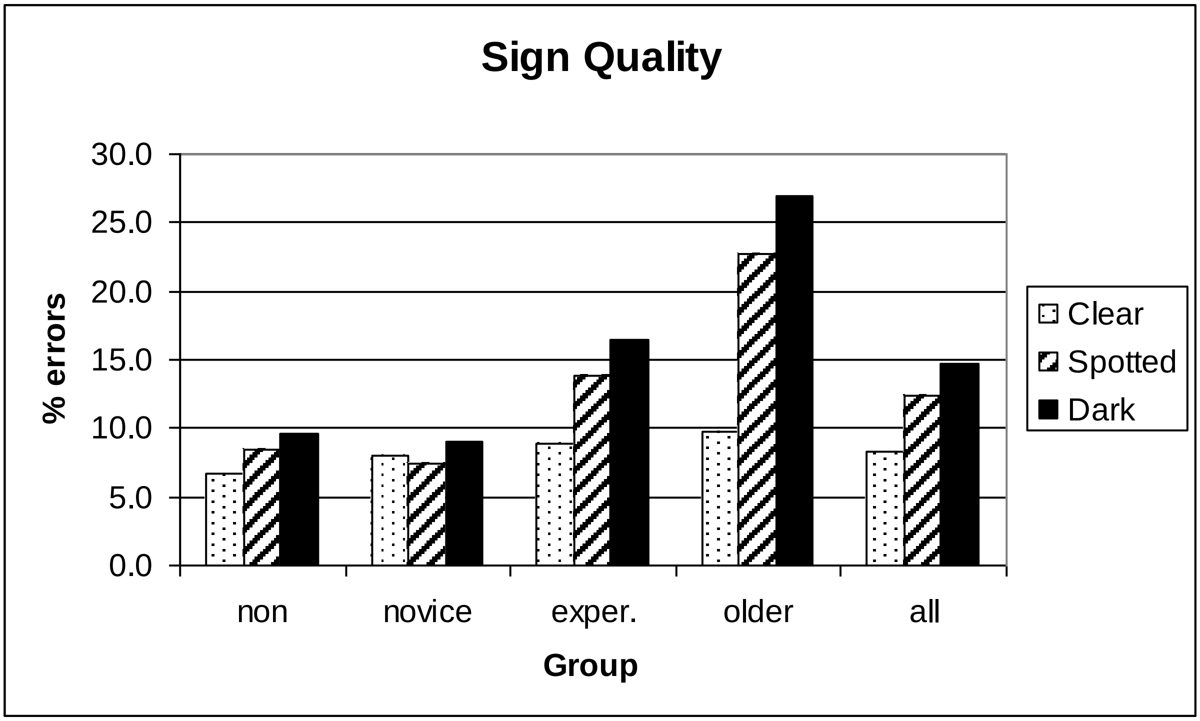
Sign Quality
30.0
25.0
20.0
15.0
10.0
5.0
0.0
% errors
non
novice
exper.
older
all
Group
Clear
Spotted
Dark

Figure 2.

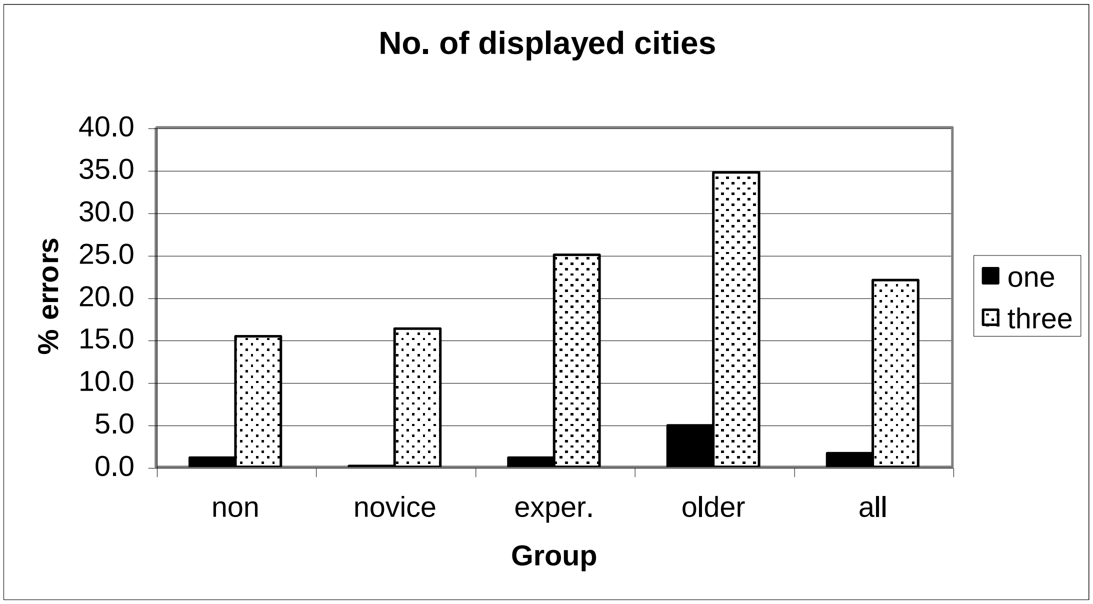
No. of displayed cities
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
% errors
non
novice
exper.
older
all
Group
one
three

Figure 3.

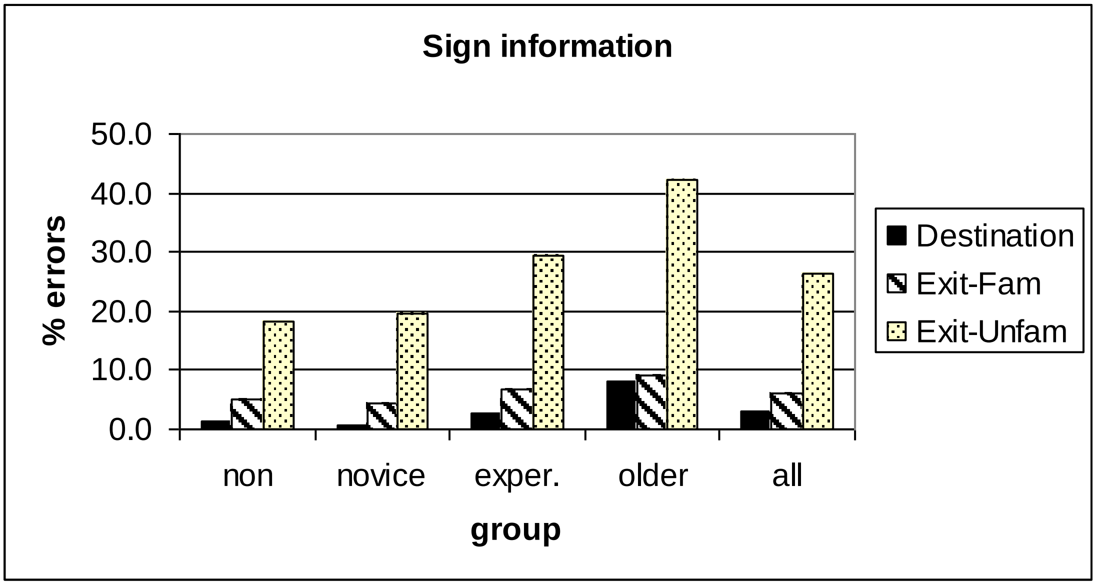
Sign information
50.0
40.0
30.0
20.0
10.0
0.0
% errors
non
novice
exper.
older
all
group
Destination
Exit-Fam
Exit-Unfam

Figure 4.

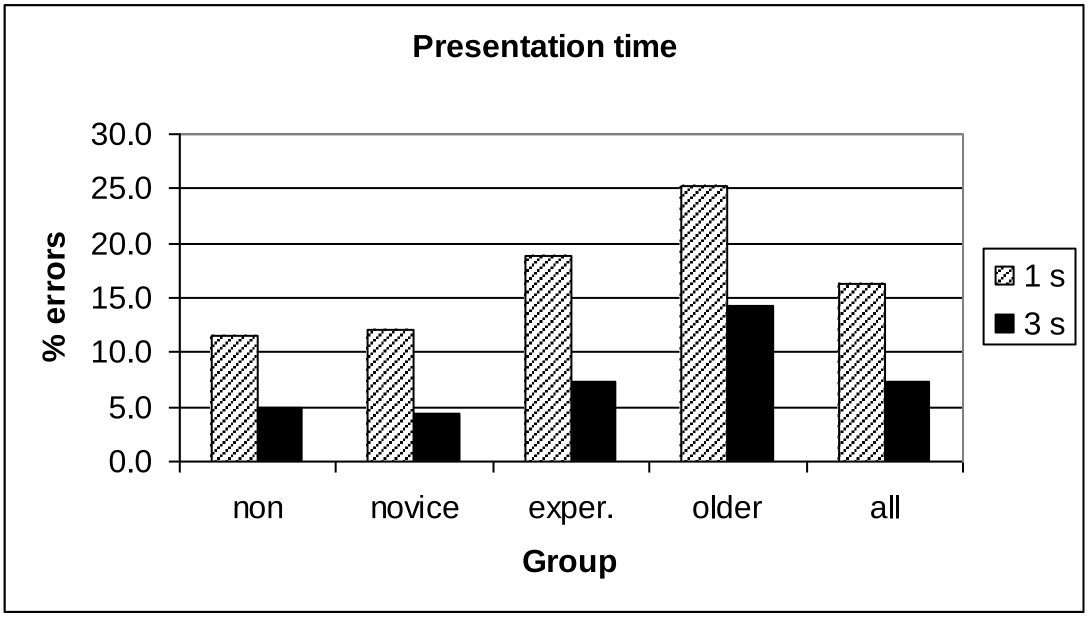
Presentation time
% errors
30.0
25.0
20.0
15.0
10.0
5.0
0.0
non
novice
exper.
older
all
Group
1 s
3 s

Figure 5.

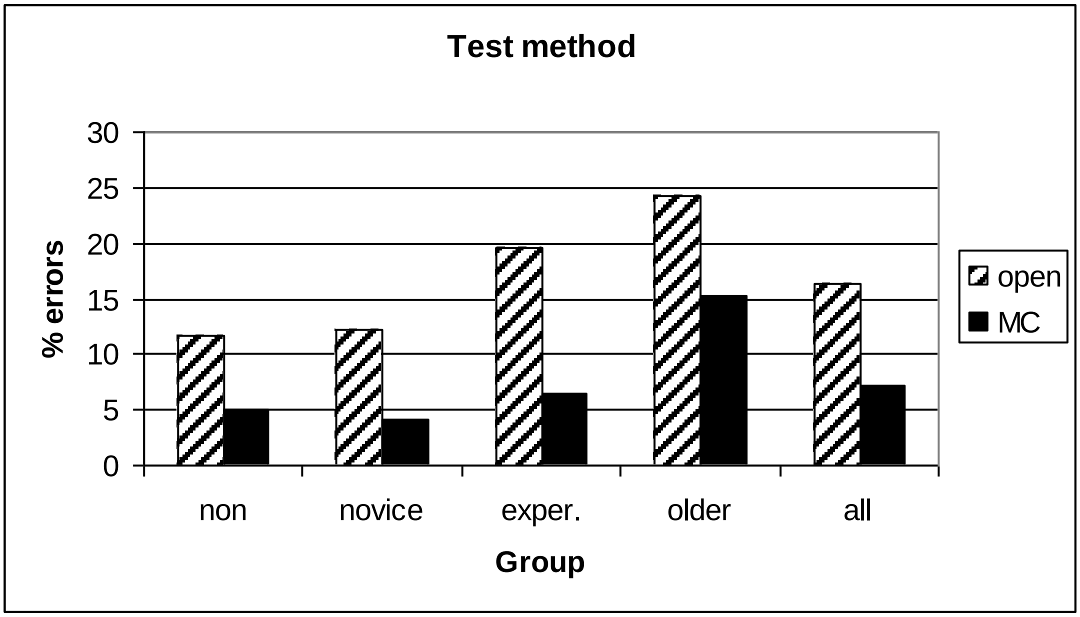
Test method
% errors
30
25
20
15
10
5
0
non
novice
exper.
older
all
Group
open
MC

Figure 6.