SELF-ASSESSMENT, QUESTIONNAIRES AND MEMORY TESTS

Chantal COMBE-PANGAUD
Lyon 2 University – CRIS/LEACM – EA 647 (France)
chantal.pangaud@free.fr Tel. (06) 70 58 27 08
Armelle JACQUET-ANDRIEU,
Paris Est University – LabInfo – UMR CNRS 8049 (France)
Armelle.jacquet@univ-paris-est.fr Tel. +331 60 95 77 19 / +336 85 75 68 89

Abstract

This paper concerns the results of a memorizing test (visual recognition and hearing recall) during an experiment of simulated driving in a magneto-encephalographic environment. The memory tests (questionnaires) are adapted to the measurement of subjective self-assessment: "tiredness" and "facility" perceptions, corralled to simple (ST) and double task (ST/DT: hearing a radio broadcast). Each subject has to answer questions and perform these tests in order to determine the psychological effect of a DT (paying attention to radio broadcasting), during his attentive driving and afferent decision-making (18 driving sessions). A debriefing questionnaire is proposed to evaluate the subject's awareness on DT effect.

Key-words

Memory test – Memorization – Visual recognition – Auditory recall – Simple Task (ST) & Double Task (DT) – Attention – Decision-making – Self-assessment – Questionnaires – Subjective perception – MEG –

1. Introduction

This presentation concerns a program of the *Agence Nationale de la Recherche* (ANR)[1] and, especially, the results off an exploration of double task effects on attention and decision making during a simulated driving experiment. These effects were studied in three different fields: Central Nervous System (CNS), by means of magnetoencephalography (MEG[2]), Autonomic Nervous System (ANS) by means of sensors, especially electrodermal activity[3], and psychological field by means of self-assessments, tests and questionnaires).

2. Material and method

The protocol is applied to thirty subjects, aged between 21 and 30 years, for an experiment consisting of a simulated driving task in a magneto-encephalography environment. There are two conditions: in the first one (simple task, ST), the subject has to drive and respect road signs (stop at orange traffic lights, use direction indicators depending on road signs) while listening to the radio. In the second condition (double task, DT), the subject has to pay attention to the radio in order to answer three questions asked before the beginning of each driving scenario. These questions are being asked again at the end of each scenario.

---

[1] A Program of the *Agence Nationale de la Recherche* (ANR) / National Research Agency: ADEME n° 0566C76, ANR-05-PDIT-004-02. PREDIT GO 4 Program: "*Technologies pour la Santé* / Technology for Public Health".
[2] MEG: Non-invasive and high performance system: 275 Channels; spatial resolution = mm; temporal resolution = ms). Experimentation conducted in the CERMEP, U INSERM 280 (Bron: France).
[3] A study of human electrodermal activity: INSA – Lyon (France) ; MMB (Micro Capteurs biomédicaux), Laboratoire de Physique de la Matière, UMR CNRS 5511.

In order to respect MEG constraints, each type of *stimuli* (traffic lights and direction road signs) is replicated 50 times in each condition. So, all subjects have to perform 18 *scenarii*, 9 in ST and 9 in DT.

> NB: for six subjects, the experimental instructions have been slightly modified. Instead of DT sessions, they were asked to drive as if they were in 9 different conditions, carrying out presumed mechanisms of appropriation like representation, identification and desire. For example, « You are driving towards the station to catch a train and you are very late », « You are driving just like your favourite hero in a detective movie and you are in pursuit of a criminal »…

3. Psychological measurements and problematic

During the experiment, the subject has to answer questions and perform tests in order to determine the psychological effect of the DT on attention and decision making while driving.
First, questions on the subject's general state (nervousness, tiredness, motivation, well-being) were asked at the beginning and ending of the experiment, using a semantic scale.
Secondly, after each driving session, the subject has to assess his tiredness degree and the task difficulty on four-point scales. These questions have been introduced to assess the subjective awareness on the DT effect, as a metacognitve knowledge or judgment (J.H. Flavell & H.M. Wellman, 1977[1]; Noël, 1997[2]; Nelson & Narens, 1990[32]).

Ultimately, at the time of debriefing, two memory tests were displayed. The first one concerned the visual environment of driving. The subject has to tell, for each of 42 objects, if he is sure to have seen it or not, during the driving task (a third alternative admits "*don't know*" and "*not sure that he has seen it or not*"). Among these 42 objects, 14 were frequent (present in each session), 14 were totally new, i.e. absent, and 14 were "rare" (present in only one of the 18 driving sessions). Our hypothesis is that rare objects encountered in ST session will be better recognized than those encountered in DT session.

The other memory test concerns the content of broadcastings (a question for each of the 16 broadcastings listened to, during the experiment). The hypothesis is that broadcastings listened to during the DT condition will be better memorized than those listened to during the ST condition, since the subject has paid better attention to them because of the previous screening for three answers while driving.

Lastly, a debriefing questionnaire has been proposed after these tests in order to assess the subject's awareness on DT psychological effect on attention and decision making while driving.

4. Results

4.1. Performance in short term memory questionnaire about auditory broadcasting during experimental session

Memorizing score in DT sessions has been computed by giving 1 point per correct answer (3 per session); so in each driving session in DT condition, the maximum score is 3 points and the maximum total score is 27. The performance is the total correct answers related to maximum. This score has been corrected by several adjustments (it is not necessary to develop this point here).
Responses were accurate in 77% of cases, showing that subjects observed the DT instructions. Individual differences were however important (sd = 0.12; minimal score = 0.37; maximal score = 0.95), suggesting that there are differences in the ability to manage the orientation of attention on broadcastings, while performing a driving task. In section 5 below, we can see that for certain subjects, the driving task is considered as the principal one while for others, it is the

memorizing task. These results may then come from the relative importance attributed by subjects at each task, despite the instructions advising the subject to concentrate on the driving task (signal respect).

4.2. General Self-assessments

Assessments vary significantly on 4 of the 5 scales. The experiment has not got any effect on subjective nervousness. However, it generates "*tiredness*", "*weariness*" and a "motivation" decline. Subjects feel though more at ease at the end of the experiment than at the beginning.

DPsy-TABLE. 1 : Scales used

a	Nervous	1	2	3	4	Relaxed
b	Tired	1	2	3	4	Rested
c	Weary	1	2	3	4	Enthusiastic
d	Not motivated	1	2	3	4	Motivated
e	Ill at ease	1	2	3	4	At ease

DPsy -TABLE. 2 : Experiment effect on subjective self-assessments. Comparison between «before/after examination» assessments

Scale	Values	Beginning	Ending	difference	sign
a	Mean Std-dev.	3.43 0.73	3.53 0.56	+0.10 t=0.71	Non sign.
b	Mean Std-dev.	3.27 0.86	2.48 0.98	-0.78 t=-4.23	P<.0005
c	Mean Std-dev.	3.53 0.63	2.97 0.89	-0.57 t=-3.20	P<.0025
d	Mean Std-dev.	3.73 0.52	3.27 0.74	-0.47 t=-3.50	P<.001
e	Mean Std-dev.	3.40 0.72	3.67 0.61	0.27 t=+2.11	P<.025

These results show, as it is confirmed by debriefing responses (see Section 3.5) and other evidences (see Section 3.3), that the experiment is quite lengthy, boring and tiring.

4.3. "Task difficulty" and "tiredness" assessments after each driving session

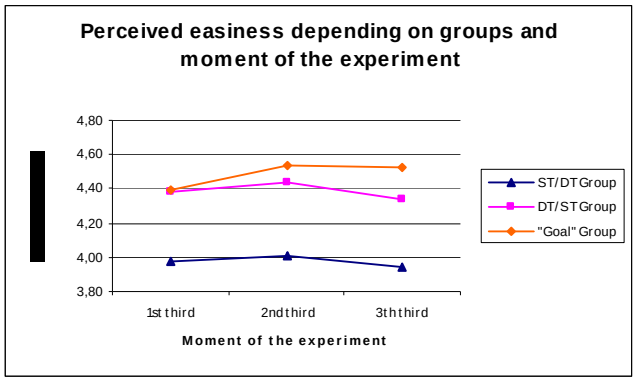
4.3.1. Experience (time) effect on assessments (4-point scale)

The purpose is to check if subjective assessments of "*difficulty perception*" and "*tiredness perception*", measured after each of the 18 driving sessions, evolves over time, index of an increasing menta- load. The set of judgments was divided in three parts, each comprising 6 consecutive *scenarii*; among these 6 *scenarii*, 3 were performed in ST and 3 in DT because of the alternation of conditions. For each subject, the mean is computed for each cluster of *scenarii*.

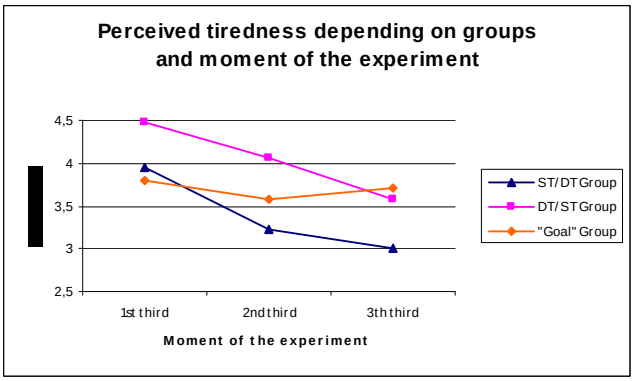
An ANOVA is computed on each variable for all subjects[4] in the same population. The "*task easiness / difficulty*" perception does not evolve over time ($F = 0.18$, ns), while "*tiredness perception*" is increasing with time ($F = 5.33$, $p < .01$). Possible differences between groups are not considered here.

[4] The Three groups of subjects are: 1. subjects beginning the experiment with a ST, 2. subjects beginning the experiment with a DT, 3. subjects for witch instructions were modified (GOALS).

DPsy-GRAPH. 1 : "Easiness / Difficulty perception"



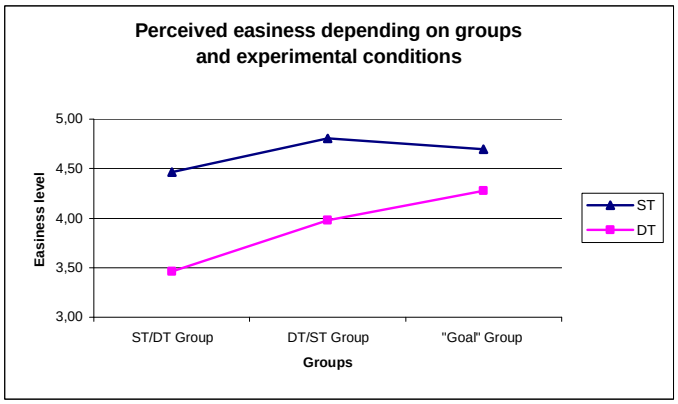
DPsy- GRAPH. 2 : "Tiredness perception"



4.3.2. Double task effect (DTE) on assessments (4-point scale)

a) Easiness perception

DPsy- GRAPH. 3 : "*Easiness / Difficulty perception*"



For each subject and condition (ST/DT), the overage of "*easiness*" evaluations is computed (5-points scale with 1 = difficult, 5 = easy); so, there are two values to compare per subject. If the subject feels that ST is easier to perform, the mean on ST may be greater than the mean on DT.

The comparison is performed by a paired two-sample t-test. The t-value on all subjects is 9.08 (df = 29, p<.0005), meaning that the subjects feel more difficulties in performing *scenarii* under DT instructions. Results are similar within each group, those beginning by a ST and those beginning by a DT. However, the 6 subjects from the last group (GOAL) show a smaller effect.
The added task in this group (drive as if they were in a variety of possible realistic conditions) seems easier than the DT consisting of searching for three answers about the radio broadcasting

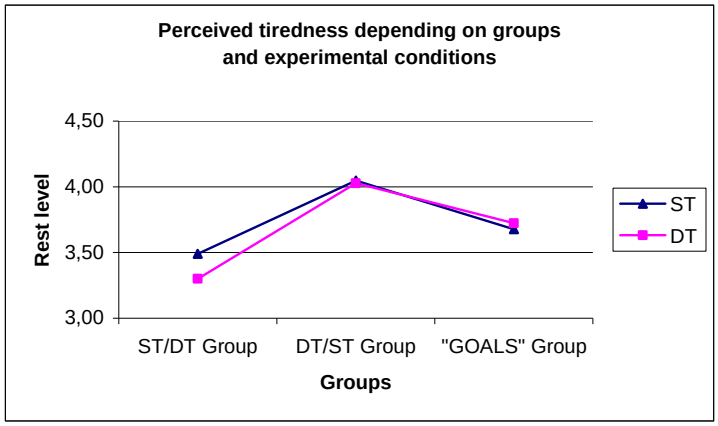
b) Tiredness *perception*
A similar analysis is performed on the « *tiredness* » assessment, noticed after each driving session (5-point scale with 1 = tired, 5 = rested). As previously, there are two values per subject, one for the 9 ST conditions and one for the 9 DT conditions.

The t-value on all 30 subjects is 2.12 (df = 29, p<.025): subjects feel more tired after a DT condition.

However, this result arises from subjects that began by a ST condition (*cf.* DPsy- GRAPH. 4). Comparisons within each group confirm that the effect of DT on the "*tiredness perception*" is only present in this group (t(10) = 3.41, p < .005). The differences are not significant in the two other groups:

- Subjects that began with a DT condition: t (12) = 0.56, ns
- Subjects in the "GOAL" condition t (5) = – 0.85 n

DPsy : GRAPH 4 : "*Tiredness perception*"



Then, the tiredness assessment applies better to the perception of the experiment length while the task easiness assessment allows the distinction between mental-load arising from DT *versus* ST. This result confirms the efficacity of easiness judgment to reflect mental representations (T.M. Ostrom & K.M. Gannon, 1996[4]; A. Jouandeau & C. Combe-Pangaud, 1999[5]; C. Combe-Pangaud, R. artin & A. Jouandeau, 2001[6]).

4.4. Long term memory results

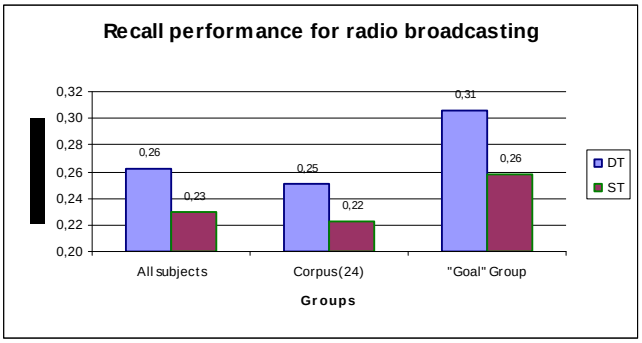
4.4.1. Double task effect on LTM for recall broadcasting content

Our hypothesis deals with a better long term recall of broadcastings content when the subjects listened to them in DT condition, because of more attention required to respond to the three questions asked before the driving session.

The score was corrected by removing answers (correct or not) for the broadcastings which were listened to twice or more during the experiment (due to experimental procedure). One point was attributed per correct answer and partial responses received ¼, ½, $^{1}/_{3}$ or $^{2}/_{3}$ depending on the degree of accuracy.

The total memorizing score is the proportion of correct answers. For the overall subjects (30), the mean score is of 0.25 (sd=0.12), that is a quarter of maximum score.
The result of interest is the score difference between the two encoding conditions (ST *versus* DT;

DPsy-GRAPH 5 : LTM Test (DT *versus* ST)



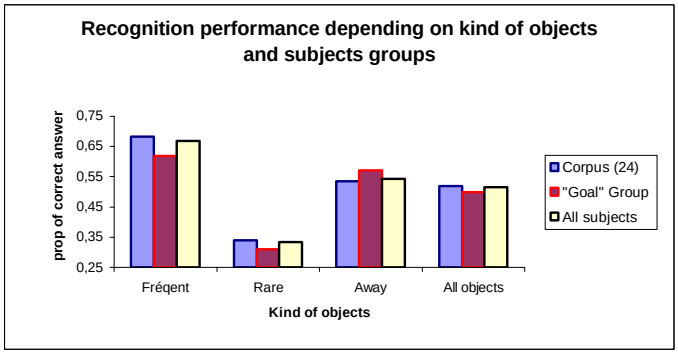
Data tend to confirm the hypothesis (DT > ST) but the difference is not significant (t (29) = 0.84, ns). There is no difference between the two groups of subjects despite an apparent better score in the "GOAL" group in the two experimental conditions (ST: t(28) = – 0.50, ns ; DT : t(28)= – 0.73, ns) ; caution is however necessary because of the small number of subjects.

4.4.2. Double task effect (DTE) on LTM for object recognition

As for radio contents, subjects' memories were tested for objects of the driving simulation environment by means of a recognition test.

For each of 42 objects, subject had to tell if "he was sure to have seen it", "he was sure to have not seen it", "the objects may or may not have been encountered". Subjects were told than the third option was not highly desirable as objects were or were not really present in the scenes.

DPsy-GRAPH. 6 : Recognition performance



Among the 42 objects used for test, there were:

- 14 objects frequently encountered during the experiment
- 14 objects appeared in one scenario
- 14 new objects, never seen

On overall 30 subjects, proportion of correct answers is 0.51 (*cf*. DPsy-GRAPH 6).
This graph shows that frequent objects are correctly recognized. New objects are also pretty well recognized as such.

Recognition for «rare objects» obtains the worse score: they are especially recognized as «absent». We can notice that the mean recognition score is near chance level i.e. 1/3.

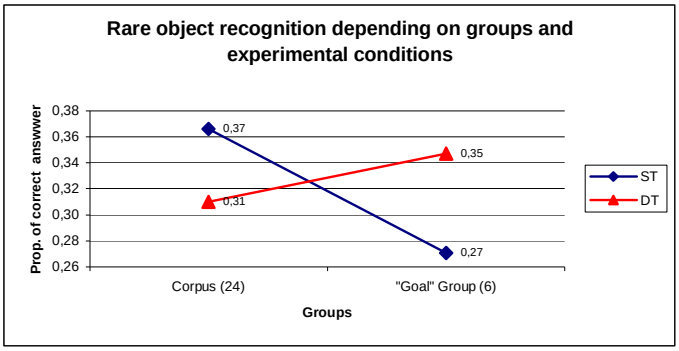
Differences between groups are small but the comparison is difficult because of the unevenness of group sizes.

Our initial hypothesis about rare objects recognition is as follows:

Rare objects might be better recognized when encoded under ST conditions. In the DT conditions, attention will mainly concern the auditory modality (at the expense of visual canal).

The two groups seem to have a different behaviour concerning LTM of rare objects (GRAPH. 7), although respective means are not very different (floor effect?)

DPsy : GRAPH. 7 : Rare objects recognition



The following analysis concerns the 24 subjects from the main corpus.

A paired two-sample t-test shows a tendency ($t(23) = 1.53$, $p < 10$ unilateral test) to difference between the recognition performance of rare objects encountered in ST versus DT, which confirms our hypothesis : recognition seems better when unique objects have been presented in a ST condition.

Data obtained among the six subjects from the “GOAL” group appear as surprising since the effect seems reversed: a better recognition of rare objects under DT conditions by a perhaps better attention allocated to environment under these instructions. We may be cautious about this dataset however because of small sample size and floor effect in performance.

4.5. Debriefing questionnaire

After the experiment, subjects had to give their level of agreement for a set of propositions (first-person written sentences) using a four-level rating scale (1: complete disagreement; 4: complete agreement). These propositions were constructed in order to check the awareness of manipulated variables, particularly the effect of DT on attention and decision making.
The participants' awareness is proved by responses (for detail, see appendix). Subjects say that DT conditions require more attention. We can also note that scenarios achieved in DT seem shorter, less monotonous, and more difficult, which confirms data on easiness/difficulty assessments. Opinions are mixed concerning pleasure and safety associated with DT *versus* ST *scenarii*.

5. Synthesis/Conclusion

On the whole, the experiment was relatively gruelling for the subjects, despite their determination and pleasure to participate. General subjective assessments, measured before / after examination, and perceived tiredness, measured after each test: control of this finding.

The experiment seems to have been less tiresome for the 6 subjects from the “GOAL” Group, but the instructions they received were not really DT ones.
Environmental conditions and timing of the experiment (3 hours) are not easily transposable to more sensitive populations like senior people.

Caution is necessary in relation to data obtained in the “GOAL” group because of important modifications in instructions, although they are of great interest.

As regards to DT effect, tests realised under ST instructions seem subjectively easier than under DT; discrimination on «perceived easiness» between ST and DT2 (GOAL conditions) is less obvious.
“Perceived easiness” is a better index of DT effect than “perceived tiredness”, which applies most to duration effects and painful aspects of the experiment.

There is no obvious effect of DT on LTM performance for broadcasting contents; objects encountered once during the experiment tend to be more easily recognized if they belonged to a ST session *versus* DT session

Awareness of DT effect exists; the “GOAL” condition (DT2) seems not to divide drivers' attention as much as the condition of listening to an audio message in order to answer three different questions (DT1). It might be, because in the first case, the attributed goal is part of the main driving task while the listening / memorizing task is dissociated from driving (attention is divided on a larger cognitive spectrum).

Good short-term memorizing performance in each session of driving shows the involvement of subjects in the DT; the comparison with results obtained from the 6 subjects of the “GOAL” Group on surprise memory tests confirms this issue.

APPENDIX

Responses on Debriefing Questionnaire

Responses 1 & 2: "disagreement" – Responses 3 & 4 "agreement"

"Constancy of the attention all through the task"
Divided opinions : 43% of the subjects tend to think that attention has not been constant
"Attention falling during the experiment"
Almost 2/3 of the subjects (19/30) say that their attention fell during the experiment (responses 3 and 4)
"Surprise due to certain events"
Certain events have surprised most of the subjects (80%): they mention the obstacles introduced in certain *scenarii*.
"Giving oneself some goals"
Many subjects (70%) say they gave themselves some goals : - Remembering the answers to the questions : mentioned 10 times - Driving well, obeying the orders : mentioned 9 times - being realistic : mentioned 2 times - paying attention to the pedestrians : mentioned 1 time - being careful to the unexpected : mentioned 1 time - taking a rest when no question : mentioned 1 time - being involved : mentioned 1 time - better managing certain turnings while in a hurry : being involved : mentioned 1 time The responses of the subjects obey the orders of the experimenter and they really involved themselves in the DT1 (listening).
"Feeling like giving up in the midfe of the experiment"
Only 6/30 of the subjects mentioned a tiredness due to the length, tiredness, lack of comfort, movements.
"Falling of motivation during the experiment"
Although very few recognized it (7/30 tick the responses 3 and 4), the evaluations on a scale between the beginning and ending of the experiment clearly show a true falling of motivation, that one can easily understand.
"Trying to go as fast as possible"
- One can see a clear distinction between the two groups: main corpus alternating with STT/DT1 *versus* Group of 6 subjects (ST/DT2). - In the main corpus, the subjects tend to say that they have not tried to go as fast as possible. : 70% of the responses rather show a disagreement between their behaviour and that assertion. - In the group "GOAL" (DT2), 50% of the subjects (3) acknowledge the fact that they were following the objective of rapidity: it is linked with the aspect of emergency ("driving like a pilot of racing", "having to rush someone to hospital").
"Increase of tiredness during the experiment"
Divided opinions: 16/30 of the subjects mention an increase of their tiredness. The two groups seem to differentiate themselves from each other: 2/6 of the subjects of the group "GOAL" are feeling signs of tiredness (33% *versus* 58% in the main corpus). The menta- load seems different for the DT2: they feel much more motivated.
"Influence of the environment"
The environment has not really got any emphasis on the subjects (≅ 90% of responses 1 or 2 on the scale).
"Easiness to obey the orders"
90% of the subjects say the orders are easy to obey (80% of the responses 4)
"Probability of the events"
72% of the subjects consider that the events are predictable: a dozen of them have noticed that the traffic lights stays green when there is a change of direction and that they often turn red on a straight line.
"Locating the regularities"
67% agree to that *item* (*cf*. previous *item*) - Duration and starting of the traffic lights: 9 times - Same rounds: 5 times - Scenery, cars stopped, pedestrians: 3 times - When one follows a vehicle, the traffic lights turn red: 1 time
"Simplicity of the driving of the simulator"
On the whole, the driving is said to be simple, but we can note some interesting remarks: - Problems of direction: 4 times - Delay time: 2 times - The indicator is a bit far: 1 time

- Condition far from reality:	1 time
- Dead angle:	1 time
"I was expecting something different"	
Several subjects (8) say they had been expecting something different:	
- More unexpected events, traps:	4 times
- More realism:	3 times
- A more diversified environment:	1 time
"The simulation is realistic"	
Divided opinions: half of the subjects choose the responses 1 or 2.	
- Little traffic, bustle in the streets:	5 times
- Motionless graphics, monotonous environment, lack of fluidity :	4 times
- Direction or pedals "not realistic":	3 times
- Breaks too slow, time of reaction badly rendered:	2 times
- Lack of angles of vision (mirrors, field of vision reduced):	2 times
- Good graphics:	2 times
- New cars:	2 times
- Vehicles easy to handle:	1 time
- The car always runs as if reversing:	1 time
- The radio talks about recent events:	1 time
- Events little frequent:	1 time
- Too binary for the breaking and acceleration:	1 time
- Feeling of speed hard to feel:	1 time
- Nothing to do with real driving:	1 time
"Resemblance with a video game"	
Rather agree with that comparison (67%): a lot of subjects underline the monotony, the lack of realism and unexpected events, the absence of notion of game which allows to differentiate with a real videogame.	
"I paid a lot of attention to the broadcastings"	
The data are reversed between the two groups: the subjects of the "GOAL" Group (DT2) were not asked to listen to broadcastings: 8 subjects out of 24 (main corpus) say they did not pay attention to the broadcastings *vs* 4/6 in the other group.	
"The DT requires a lot of attention"	
One can distinguish the two groups according to DT1 (radio listening) & DT 2 (particular objective of driving). The first subjects (96%) say the task requires a lot of attention *vs* 17% for the other one. That confirms the data we have got about "perceived easiness" of the task between the two conditions.	
"The memorizing task disturbs that of driving (corpus)"	
The subjects rather disagree with that statement (14/23)	
"The driving task disturbs that of memorizing (corpus)"	
17/24 say "yes". The memorizing task seems more difficult than that of driving: less automatism so more attention requested.	
"The task of driving has been preferred to that of radio listening (corpus)"	
Divided opinions: 13/24 subjects say they have preferred the task of driving.	
"Comparison of the routes between DT and ST" (corpus)	
The routes in DT prove to be : - shorter (16/23); - less monotonous (21/24); - more difficult (18/23); which confirms the data analysing "perceived easiness"; - divided opinions regarding the pleasure (13/24 find them less comfortable); - divided opinions regarding the level of associated risk (13/24 find them more risky).	
"The broadcastings are disturbing the carrying out of the task" ("GOAL").	
Most of the subjects (5/6), say they have not been disturbed by the broadcastings, which they did not have to remember.	
"The driving is disturbed during the simulation exercise (DT2)" ("GOAL")	
Most subjects agree with that (4/6): all of them say they focused on the orders of simulation exercise ("GOAL").	
"Having preferred the task of driving for assigned aim" ("GOAL")	
Most subjects say "no" (5/6): they have preferred the "GOAL" to follow, to the extent of sometimes not following the driving code. The distances with "GOAL" (DT2) appear: - to be shorter (5/6) - less monotonous (6/6) - more difficult (5/6) - less pleasant (5/6) - more risky (6/6). On the whole, the responses of those subjective assessments are coherent with those collected in the main corpus of subjects but are much more contrasting.	

References

[1] Flavell, J.H. & Wellman, H.M. (1977). *Metamemory*, in R.V. Kail & J.W. Hagen (Eds), Perspectives on the development of memory and cognition, Hillsdale, N.J.: Lawrence Erlbaum Associates.
[2] Noël, B. (1977). *La métacognition*. Bruxelles : De Boeck Université.
[3] Nelson, T.O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. in G. Bower (Ed), *The Psychology of Learning and Motivation*, vol. 26, pp. 155-160.
[4] Ostrom, T.M. & Gannon, K.M. (1996). *Exemplar generation: Assessing how respondent give meaning to rating scales*, in N. Schwarz & S. Sudman (Eds), Answering questions: Methodology for determining cognitive and communicative processes in survey research, San Fransisco: Jossey-Bass Inc. Publishers, pp. 293-318.
[5] Jouandeau, A. & Combe-Pangaud, C. (1999). Baromètre de satisfaction. Evaluation des représentations mentales des quantificateurs d'appréciation, *Les Cahiers de la Recherche Scientifique en Qualité*, 6, pp. 1-18.
[6] Combe-Pangaud, C., Martin, R. & Jouandeau, A. (2001). *Le sentiment de facilité, indicateur de la variabilité interindividuelle des réponses*, in A. Flieller & al., *Questions de Psychologie Différentielle*, Rennes, Presses Universitaires de Rennes, pp. 81-86.