Manipulating Stress and Cognitive Load in Conversational Interactions with a Multimodal System for Crisis Management Support

Andreea Niculescu, Yujia Cao, and Anton Nijholt

Human Media Interaction, University of Twente
P.O. Box 217, 7500 AE, Enschede, The Netherlands
{niculescuai,y.cao,a.nijholt}@utwente.nl

Abstract. The quality assessment of multimodal conversational interfaces is influenced by many factors. Stress and cognitive load are two of most important. In the literature, these two factors are considered as being related and accordingly summarized under the single concept of 'cognitive demand'. However, our assumption is that even if they are related, these two factors can still occur independently. Therefore, it is essential to control their levels during the interaction in order to determine the impact that each factor has on the perceived conversational quality. In this paper we present preliminary experiments in which we tried to achieve a factor separation by inducing alternating low/high levels of both stress and cognitive load. The stress/cognitive load levels were manipulated by varying task difficulty, information presentation and time pressure. Physiological measurements, performance metrics, as well as subjective reports were deployed to validate the induced stress and cognitive load levels. Results showed that our manipulations were successful for the cognitive load and partly for the stress. The levels of both factors were better indicated by subjective reports and performance metrics than by physiological measurements.

Keywords: Multimodal interfaces, verbal communication, stress, cognitive load, physiological measurements, qualitative evaluations.

1 Introduction

Nowadays, multimodal conversational interfaces enable users to communicate with computer systems using a wide range of input/output modalities, such as speech, text, touch, etc. Therefore, there is a growing need not only to find reliable evaluation methods for such interfaces but also to determine which factors have the highest impact on their quality assessment. The literature in the field mentions the stress and cognitive load experienced by users while interacting with an interface among the most important influence factors [1].

Cognitive load is often described as the degree of concentration required for a person to solve problems or to complete tasks in a given time [2]. The term, referred in the literature as 'cognitive load'[2], 'cognitive or mental effort' [3], 'cognitive factor' [4] is often associated with the factor 'stress'. Stress represents a psychological response

A. Esposito et al. (Eds.): COST 2102 Int. Training School 2009, LNCS 5967, pp. 134–147, 2010.

state to a perceived threat or task demand and is in general signalized by specific emotions such as frustration, anxiety and tenseness [5]. Stress is caused by an existing stress-causing agent that can be a physical-environmental stressor (e.g. noisy environment) or psychological stressor (e.g. work overload). In the literature, both factors are often summarized and measured under the global concept of ‘cognitive demand’ [6].

There are no doubts that cognitive load and stress are related. When the load reaches a certain level of demand - people unconsciously appraise their abilities to meet the challenge: only if the situation is considered as exceeding the available resources then stress would appear. This appraisal theory was formulated by Lazarus et al. [7] and explains why a tense situation might be perceived as stressful by one person, but not by another.

Based on this theory, our assumption is that, even if related, these two factors can occur independently, i.e. there is no compulsory relationship between them. Also, the stress might be produced by other variables that do not directly relate to cognitive load, such as background noise, frequent misunderstandings, or increased interaction speed. Thus, these two factors might have a different impact on the perceived conversational quality and consequently, they should be identified and measured separately.

Hence, we propose an experiment intended to determine the circumstances in which low/high levels of both stress and cognitive load are alternatively achieved while interacting with a multimodal conversational system for crisis management support. Once such circumstances are identified they will be incorporated in the design of a further study, investigating the impact of stress and cognitive load on the perceived conversational quality.

2 Experiment Set-Up

2.1 Scenario and Trials Design

Crises are situations in which people experience high levels of stress and cognitive load. Therefore, they offer a perfect test environment for our experiments. Accordingly we designed four scenario trials based on a common typical crisis situation: an explosion occurred in a chemical research lab setting an entire floor on fire; a crisis manager in charged of the situation has to take essential, life-critical decisions based on the information received from the system.

Each trial consists of three stages: at first, a virtual assistant, representing the system, presents the current crisis situation using narrative, assisting photos, maps and text; next the crisis manager has to find, using an interactive map, addresses to which rescue workers, fire fighters will be sent or wounded victims will be delivered; alternatively he has to memorize important event facts and insert them in a crisis report; finally, the crisis manager gets chemical description sheets to identify dangerous chemicals that have to be immediately removed by firemen in order to avoid further damages.

2.2 System Design

A first multimodal system prototype was developed using the CSLU toolkit[1]. The prototype, currently under development, provides detailed information about the crisis

[1] The toolkit can be downloaded at: http://cslu.cse.ogi.edu/toolkit/

event, such as event description, geographical maps, available rescue resources and estimated number of victims. The system has attached an embodied conversational agent with text-to-speech and speech recognition capabilities. Test users can interact with the system using speech or mouse clicking and receive information in the form of text, speech, images or videos.

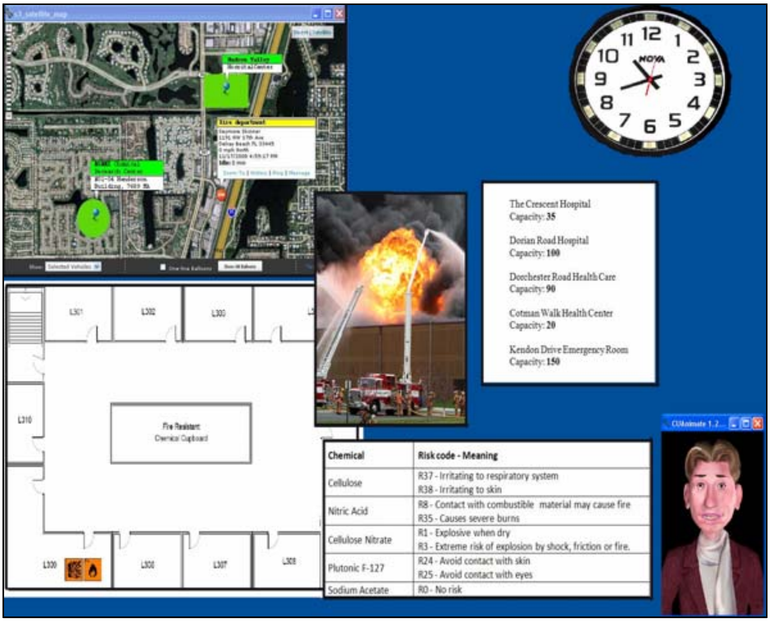

Fig. 1. System screen shot

2.3 Factor Manipulation

The trials were similar but aiming to realize different combinations of low/high stress and low/high cognitive load conditions in a 2x2 factor matrix.

In order to manipulate the stress level we used a combination of six different parameters such as background noise, speech speed, speech length, time limitation, simulated recognition mistakes and dramatic event description. In the low stress conditions, the virtual assistant presents calmly the crisis event using a clear voice with normal speed. He describes the situation as being under control; users are not urged to speed up their performance. In the high stress conditions however, noises (e.g. white noise, ambulance sound) are played in the background in order to induce stress [8]. The agent talks faster, using short sentences and an urgent tone. The crisis situation is described to be dramatic; users are put under time stress by being constantly reminded to make quick decisions; a simulated speech recognition mistake was also built into one of the scenarios.

For the cognitive load manipulation we used two parameters: task complexity and presentation format. Task complexity variations were put into effect for address identification: in the scenarios with low cognitive load the users had to locate given addresses on a map by clicking on the street names; in the scenarios with higher load the users were required not only to identify but also to select the optimal address, according to several factors that needed careful analysis (e.g. hospital capacity, distance to the chemical lab, number of victims). Variations in the presentation format were chosen for the chemical selection task. 'Well'-designed and 'badly'-designed information sheets were applied to achieve low and high cognitive load conditions respectively. Both sheets use a table to present the chemicals and their risk descriptions. The difference between the 'well'- and 'badly'-designed sheet lie in the way the information is spatially organized: the 'well'-designed sheet provides integrated chemicals and risk descriptions in a natural 'row-by-row' sequence facilitating the users 'scan'-reading; in contrast, the 'badly'-designed sheets provide numerical codes that links the chemicals to their corresponding risk descriptions summarized outside the table. As a consequence, the 'badly'-designed sheet requires additional mental effort, causing a split-resource effect and an increase of cognitive load [9].

Chemical	Risk code - Meaning
Cellulose	R36 - Irritating eyes R37 - Irritating to respiratory system R38 - Irritating to skin
Chlorine	R23 - Toxic by inhalation R50 - Very toxic to aquatic organism
Toulene	R11 - Highly flammable in heat R63 - Possible risk of harm to the unborn child R65 - May cause lung damage if swallowed
Alumina	R0 - No risk
Barium chloride	R20 - Harmful by inhalation R25 - Toxic if swallowed

a)

Chemical	Risk code
Benzyl phthalate	R50 R53 R62
Nitric Acid	R8 R35
Dimethyl formamide	R61 R24 R43 R21
Ammonia	R14 R23 R45
Benzyl peroxide	R2 R36 R6

R36 Irritating to eyes
R14 Reacts violently with water
R23 Toxic by inhalation
R45 May cause cancer
R2 Risk of explosion by fire or other sources of ignition
R61 May cause harm to the unborn child
R24 Toxic in contact with skin
R50 Very toxic to aquatic organisms
R6 Explosive with or without contact with air
R43 May cause sensitization by skin contact
R21 Harmful by inhalation and in contact with skin
R8 Contact with combustible material may cause fire
R53 May cause adverse effects in the aquatic environment
R62 Possible risk of impaired fertility
R35 Causes severe burns

b)

Fig. 2. "Well" designed (**a**) vs. "badly" (**b**) information sheets

The following table presents a summarization of the factor manipulations in the four trials.

Table 1. Cognitive load and stress manipulation per trial

Trial nr.	Trial description	Cognitive load (CL)	Stress level (S)
1.	T00= low CL/ low S	address identification=1 address decision task=no memory retrieval task=no presentation format= 'well' designed	background noise=no speech speed=normal speech length=normal time limitation= no recog. mistake=no dramatic event description=no
2.	T01= low CL/ high S	address identification=5 addresses decision task=no memory retrieval=no presentation format = 'well' designed	background noise=yes speech speed=high speech length=short time limitation= no recog. mistake=no dramatic event description=yes
3.	T11= high CL/ high S	address identification=1 addresses decision task= yes memory retrieval=no presentation format='badly' designed	background noise=yes speech speed=high speech length=short time limitation= no recog. mistake=yes dramatic event description=yes
4.	T10= high CL/ low S	address identification=no decision task= no memory retrieval=yes presentation format='badly' designed	background noise=yes speech speed=high speech length=short time limitation= no recog. mistake=yes dramatic event description=yes

3 Methods

3.1 Measurements

For our experiment we used a combination of several assessment methods such as subjective rating, physiological measurements and performance metrics.

The subjective ratings were collected after each trial using the NASA task load index (TLX) questionnaire. NASA-TLX contains six workload-related parameters: mental, physical and temporal demands, own performance, effort and frustration. The level of frustration is measured by NASA-TLX with the help of a single question addressing simultaneously five different parameters: feeling insecure, discouraged, irritated, annoyed, and stressed. We considered appropriate to split the question in five separate statements (one for each parameter) in order to get more precise results. We also replaced the term "stressed" – a key concept in our study - with the semantically related word "tense". In our analysis, we treated the concept of "tenseness" apart from frustration, since we consider that these two categories may not always be related.

We added four additional statements to the TLX questionnaire regarding the users' concentration and tiredness level, the system's easy of use, the degree of understanding between users and system.

We relate the frustration, tenseness and temporal demands (work pace) parameters as direct indicators for the stress factor and considered the other parameters as direct indicators for the cognitive load [10]. A 20 level scale was used for ratings.

As physiological measurements, we used the heart rate variability (HRV) and the galvanic skin response (GSR). According to previous studies, certain components of HVR exhibit systematic and reliable relationships with the mental demands of the task. Higher levels of cognitive work load have been associated in the frequency domain with decreased power in the 0.10 Hz band (LF – low frequency band) [11, 12]. Skin conductance response (SCR) is traditionally associated with workload and especially with arousal states accompanied by mental effort and emotions. Higher workload normally yields higher number of responses (or longer SCR intervals) [13, 14].

All trials were recorded with a video camera in order to allow the retrieval of performance metrics.

3.2 Experiment Setup

Four male test users, aged between 24 and 30, all having technical background, participated in the experiment.

After entering the lab and taking a seat, each user was asked to stay relaxed while the physiological sensors were applied. When finished, a physiological baseline was recorded for 5 minutes. Afterwards, the user received a brief introduction of the experiment and performed the trials. A short break was placed after each trial to allow test users to fill in the questionnaire and have a rest.

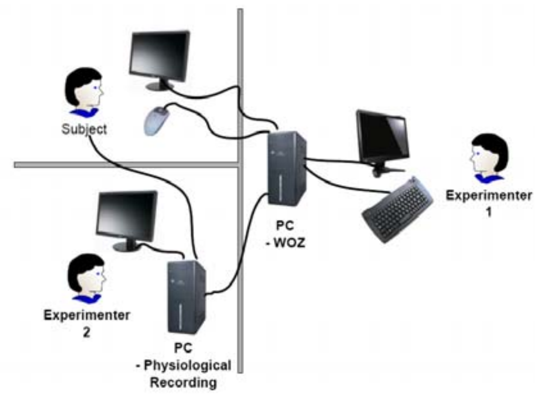

Fig. 3. Experiment set-up

The experiments were performed using the Wizard-of-Oz technique: the speech recognition module was replaced by a human operator in order to ensure a controlled interaction.

Two experimenters were involved in conducting the trials: one was in charged of the physiological measurements, the other one was performing the WOZ simulation. The experiment setup was synchronized as illustrated above (see fig.3).

4 Results

4.1 Subjective Questionnaires

The first two trials – 1 and 2 (T**0**0, T**0**1) - were designed to have a lower cognitive load level compared with the last two trials - 3 and 4 (T**1**1, T**1**0). The results[2] gathered from

[2] All results have discrete values; they are presented on a connected line only to facilitate the view.

the questionnaires confirmed that test users perceived the last two trials as being more mentally demanding (see fig.4), harder to accomplish and in general requiring a higher degree of concentration compared with the first two trials. Only user 3 indicated a low level of concentration for trial 4 (see fig.7), a fact that corresponds to his indeed low performance during this trial - the user had extremely long response times and frequent input errors.

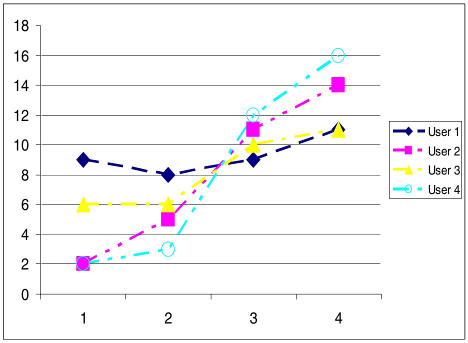

Fig. 4. Mental demand perception

The statements that the system delivered were perceived as less clear during the last two trials, and users perceived their own performance as being the worst mostly during the 3rd trial. On the other side they judged the system as being the easiest to use during 2nd trial.

Concerning the physical demand, the values were quite similar for all trials (only one user indicated higher values of physical demand for the cognitive loaded high trials).

The 4th trial was considered as being the most tiring among all trials. This trial lasted the longest as shown in the figure below (fig.5).

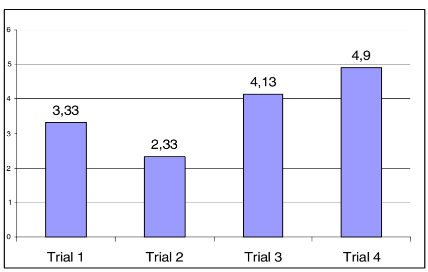

Fig. 5. Trial average completion time in minutes

Interestingly there are not many similarities between the graphs representing the users' tiredness level on one side and their concentration level on the other side (see fig. 6 and 7). We were in fact expecting users to feel more tired after completing a task requiring a high level of concentration.

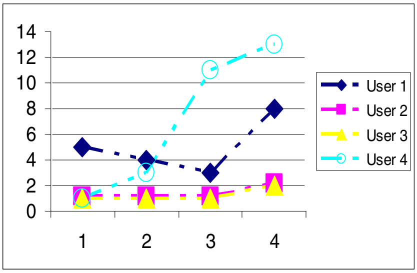

Fig. 6. Tiredness

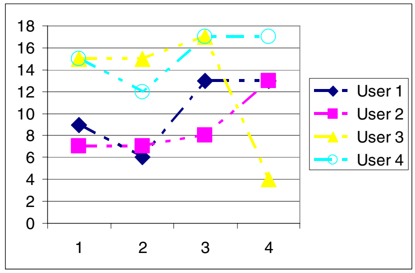

Fig. 7. Level of concentration

Regarding the stress factor, we tried to induce a lower level of stress in trials 1 and 4 (T**0**0, T1**0**) compared with a higher level in trials 2 and 3 (T0**1**, T1**1**).

The perception of temporal demand (work pace rush) was, as expected higher for trials 2 and 3.

However, the same trend could not be observed for the users' degree of frustration: participants felt more frustrated - i.e. more insecure, discouraged, irritated and annoyed - mostly during trials 1and 3 (especially trial 3, which had the highest negative values). The fact that trial 1 achieved a higher level of frustration than expected, might be explained by 'first impression' effect: during the first trial, users were dealing with an unfamiliar situation that apparently caused frustration; afterwards, they must have felt more confident with the system, and accordingly ranked trial 2 much lower in terms of frustration. There was a general up and down in the participants feeling of being insecure as shown in fig. 8.

Comparing trial 4 with trial 3 - both highly cognitive loaded - we observed that trial 4 apparently caused a much lower frustration level; this fact might confirm our assumption that a highly mentally demanding task is not necessarily accompanied by stress.

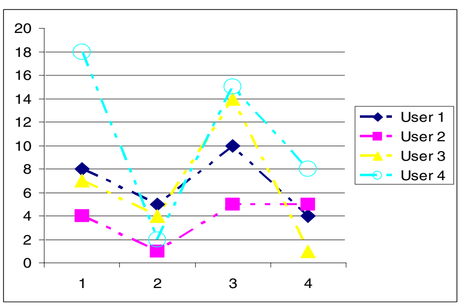

Fig. 8. Feeling insecure

The ranking of the last parameter, the feeling of being tense, showed how different the test participants perceived the trials (see fig. 9): users 1 and 2 indeed felt tenser during trial 2 compared with trial 1 and during trial 3 compared with trial 4. User 3 and 4 felt tenser only during the trial 3; trial 2 was perceived by both users as much more relaxed.

Interestingly, trial 1 was perceived as much tenser for users 3 and 4 compared with user 1 and 2. We compared the trend for this particular trial with the users' perception of being successful and we observed a certain similarity between the graphs; unfortunately, the similarity was not confirmed for the other trials, therefore we cannot make a sustainable association between tenseness and the perceived performance success.

Trial 4 was perceived by all users, except one (user 4), as being less tense when compared with trial 3.

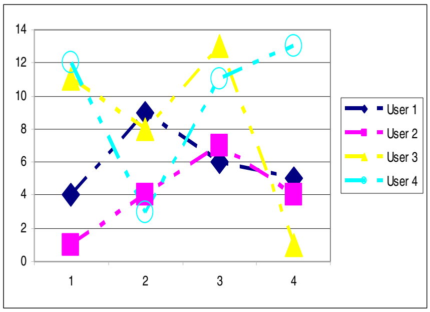

Fig. 9. Feeling tense

Clearly, despite some similarities among participants, the graphic above doesn't show a common trend. Apparently users had similar opinions about cognitively high loaded tasks, but quite different perceptions about the stress level produced while performing these tasks.

4.2 Physiological Measurements

The physiological measurements mainly showed a learning effect: the value of HR (heart rate) decreased and LF (the HRV in the frequency domain) increased trial by trial, both indicating a gradually decreasing cognitive load and stress throughout the experiment.

GSRN (number of skin responses per minute) and GSL (the tonic level of the skin conductance) showed the same effect for two of the users (users 2 and 4), as shown in fig. 10 – indicating their stress level was continually decreasing.

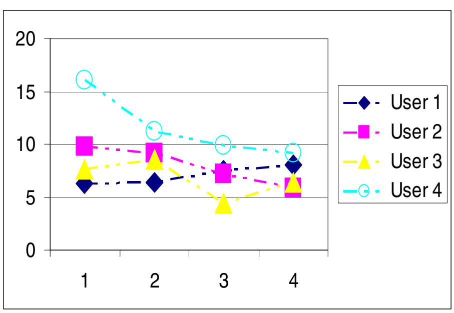

Fig. 10. GSRN values

One user's (1) stress level seemed to be caused by the cognitive load level rather than by the planed manipulation: the values showed that the user was more stressed when the task was more difficult (trial 3 and 4). The values for the last user (3) did not deliver any explainable results.

4.3 Users' Performance Metrics

We extracted from the videos the following users' performance metrics relating to cognitive loaded: the response competition time, reaction time and number of errors. For the stress metric we counted the number of words, verbal hesitations, breaks and mispronunciations.

Each trial was composed of one, two or three tasks, such as **chemical removing tasks** (in trial 1,2,3,4), **address identification tasks** (in trial 1,2,3), **decision tasks** (in trial 2), and **memory retrieval tasks** (in trial 4).

The chemical removing task consisted of three subsequent subtasks, users had to:

- 1) ask for a floor map (using speech)
- 2) localize the room with dangerous chemicals on the map (using mouse click)
- 3) find, based on the "well/badly" designed description sheets, which chemical to remove (using mouse click).

The first and the second subtasks showed a clear learning effect: test users needed on average 2.1 sec. less to ask the question and 1.2 sec. less to localize the room. The 3rd subtask concerning the information presentation format delivered more interesting results: the graphic shows that users 1, 2 and 3 took considerably more time to identify and remove the chemicals when the information sheet was badly designed. Trial 1, even if it had a well-designed sheet, achieved a much higher value than expected due to the same "first impression" effect (see fig. 11).

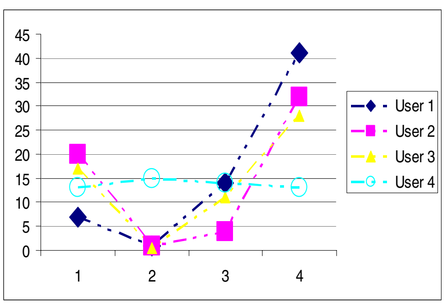

Fig. 11. Competition time for identifying dangerous chemicals

The results achieved by user 4 need to be considered separately, since the user has professional chemical expertise. His completion times are in contrast with those obtained by the other participants. Therefore, we have reasons to believe that the performance time was indeed influenced by the information presentation format and not by other parameter such as tiredness, concentration level or interaction speed.

The address identification task had two subtasks. The participants were required to

- 1) ask for a street map (using speech)
- 2) localize an address on each time different map (using mouse click).

Results showed again a clear learning effect for both subtasks, with users spending an average 16.78 sec. less each time solving the task (see fig.12). The graph is not complete for user 1, who didn't complete this task during trial 3, due to a wizard error.

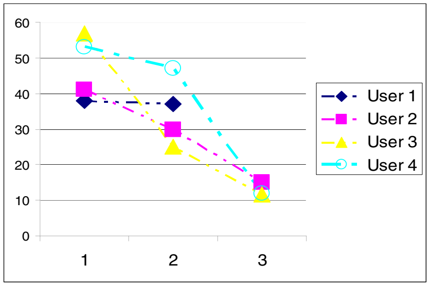

Fig. 12. Completion time for the address identification task

The performance time for the decision and memory retrieval tasks showed quite large differences among users (see tab.2). The differences seem to be related to the perceived concentration level – except for user 3, who indicated a very low concentration level despite the time he spent to solve the task. The reason was, as explained previously, the obvious lack of concentration during the trial (see fig.7).

Table 2. Decision and memory retrieval task times vs. level of concentration values

User	**Decision task -trial 3-**	**Memory retrieval –trial 4-**	**Level of concentration on 20 point scale**	
			-trial 3-	**-trial 4-**
1	36 sec.	105 sec.	13	13
2	9 sec.	90 sec.	7	13
3	52 sec.	134 sec.	17	4
4	62.s	322 sec.	17	17

According to the tasks users have to complete, they are required to interact with the system differently. Hence, we differentiate between speech responses to speech input and clicking responses to visual input. We analyzed the users' speech response time (time slot between system input and first user's reaction) and observed that the values for speech responses were higher in the first trial (in average 4.1 sec.), decreasing in the following trials 2 and 3 (1.24 sec. and respectively 0.83 sec.) before increasing again in the last trial (2.41 sec.). The decreasing trend for trial 2 and 3 might be an indication that users tried to adapt their verbal behavior to the system's speech rhythm increase.

The response time to visual tasks was higher compared with the verbal responses and did not differ much across the trials: users needed an average of 27.29 sec. for identifying an address on the map and 16.5 sec. for finding the room with chemicals.

Looking at the error distribution (see fig. 13) most of the errors were committed during the 3rd and 2nd trial – the most rushed trials. We identified 4 different types of errors such as speech errors, visual errors (clicking on a wrong target), decisions errors (making a wrong choice) and memory retrieval errors. Most of the errors were memory retrieval and decision making errors, as shown in the figure 11. The values for the error type were normalized to the corresponding number of task types (there were in total 8 speech-based tasks, 6 visual-based tasks, 4 decision making tasks and 1 memory retrieval task).

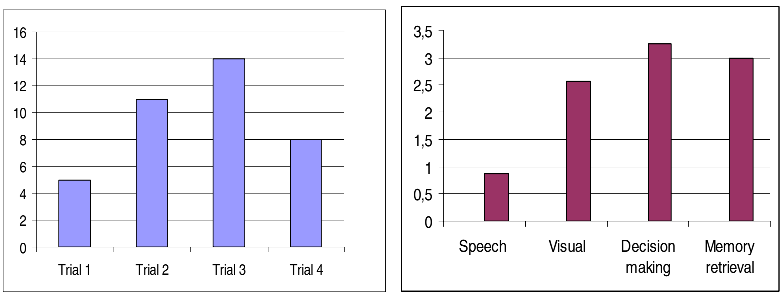

Fig. 13. Error distribution and types

An interesting observation concerns the user's speech behavior during the trials: the total amount of words and verbal breaks were on average higher during trial 1 and 4 (12.6 respectively 4.75) compared with trial 2 and 3 (8.79. 2.5). This behavior demonstrated again the users' unconscious adaptive behavior to the speech rhythm imposed by the system, reacting faster and with shorter statements during the rushed trials 2 and 3. Also other speech disfluecies, such as mispronunciations, occurred only during trial 2 and 3 which indicates a possible increase of stress during these trials.

5 Conclusion

The results showed that our cognitive load manipulation was successful: users perceived trial 3 and 4, as mentally more demanding, harder to accomplish and requiring a higher degree of concentration compared with trials 1 and 2. During trial 3 and 4, the system's statements appeared to be harder to understand and participants perceived the interaction as more difficult. Their tiredness degree was the highest mostly during the last trial - the one having the longest competition time. The information format presentation as well as the decision task complexity seemed to contribute successfully to the proposed manipulation.

The success of the stress manipulation was disturbed by 'first impression' effects: test users indeed perceived trials 2 and 3 as being more rushed, but they felt more frustrated during trials 1 and 3. The frustration degree for the mentally demanding trials 3 and 4 was in general perceived differently: trial 3 was considered as being more frustrating compared with trial 4, a fact that might support our assumption that stress does not necessarily accompany high cognitive high loaded tasks. Thus, the planed stress manipulation was successful only for trial 3. Due to an unplanned high stress level achievement during trial 1 it seems rather difficult to make assumptions about which other (planed) stressor particularly contributed to this manipulation.

Generally, the stress manipulation appeared to be relatively difficult to induce compared with the cognitive load manipulation. One reason might be the fact that stress is a highly complex phenomenon, including aspects that we did not consider in our experiment such as "first impression" effects. Another possible explanation could be that people perceive the stress very differently according to own individual dispositions [7]. These dispositions are not always influenced by people's performance success, as we might have expected: for instance our results did not show a clear relationship between the amount of errors and the perception of own performance success on one side and the users' feelings of frustration and tenseness on the other side.

Also, a less sharp stress perception on participants' side might have weakened the planned stress manipulation: the users' verbal behavior indicated more relaxed feelings during trials 1 and 4 and more stressed reactions during trials 2 and 3; these observations were not confirmed by subjective reports concerning the tenseness.

The lack of reliable objective measurement results did not help the understanding of the stress phenomenon in the experiment context. In fact, both factors – cognitive load and stress - could be better determined by subjective reports and performance metrics than by physiological measurements.

In conclusion, we consider the current experiment a good starting point for forthcoming investigations concerning the effects of stress and cognitive load on the conversational quality assessment. In the future we plan to perform similar experiments with a larger number of users, in order to gain statistical evidence for our findings. "First impression" effects will be avoided by using training sessions before starting the experiment. Also, the interpretation of physiological data can be improved by measuring particular tasks inside each trial, rather than using the whole trial (as performed in this experiment). Further, enlarging the variance in cognitive load and stress between trials might also enhance the effectiveness of physiological measurement since their sensitivity is limited to minor variations.

Acknowledgments

This research is done in the framework of the Interactive Collaborative Information System (ICIS) Project, sponsored by the Dutch Ministry of Economic Affairs, under grant nr: BSIK03024. We thank all test participants for their effort and time. A special thank to Dr. David Salamon for his help with the design of the chemical information sheets.

References

1. Moeller, S.: Quality of telephone-based spoken dialogue systems. Springer, New York (2005)
2. Oviatt, S.: Human-centered design meets cognitive load theory: designing interfaces that help people think. In: Proc. of the 14th annual ACM international conference on Multimedia, Santa Barbara (2006)

3. Love, S., Dutton, R.T., Foster, J.C., Jack, M.A., Stentiford, F.W.: Identifying salient usability attributes for automated telephone services. In: Proc. of the 3rd Conf. on Spoken Language (ICSLP), Yokohama, Japan (1994)
4. Jack, M.A., Foster, J.C., Stentiford, F.W.M.: Intelligent dialogues in automated telephone services. In: Proc. 2nd Int. Conf. on Spoken Language Processing (ICSLP), Banff, Canada (1992)
5. Patil, S.A., Hansen, J.H.L.: Speech Under Stress: Analysis, Modeling and Recognition. In: Müller, C. (ed.) Speaker Classification I: Fundamentals, Features, and Methods, pp. 108–137. Springer, Heidelberg (2007)
6. Hone, K.S., Graham, R.: Towards a tool for the subjective assessment of speech and System Interfaces (SASSI). Natural Language Engineering 6(3-4), 287–303 (2000)
7. Lazarus, R.S.: Theory based stress measurement. Psychological Inquiry 1, 3–13 (1990)
8. Kryter, K.D.: The handbook of hearing and the effects of noise: Physiology, psychology, and public health. Academic Press, New York (1994)
9. Chandler, P., Sweller, J.: Cognitive Load Theory and the Format of Instruction. Cognition and Instruction 8, 293–332 (1991)
10. Schell, K.L., Grasha, A.F.: State anxiety, performance accuracy, and work pace in a simulated pharmacy dispensing task. In: Perceptual and Motor Skills, vol. 90, pp. 547–561 (2000)
11. Scerbo, M.W., Freeman, F.G., Mikulka, P.J., Parasuraman, R., Di Nocero, F.: The efficacy of psychophysiological measures for implementing adaptive technology. TP-2001-211018 (2001)
12. Wilson, G.F., Eggemeier, F.T.: Psychophysiological assessment of workload in multi-task environments. In: Damos, D.L. (ed.) Multiple-task performance. CRC Press, Boca Raton (1991)
13. Verwey, W.B., Veltman, H.: Detecting short periods of elevated workload. A comparison of nine workload assessment techniques. Applied Experimental Psychology 2, 270–285 (1996)
14. Boucsein, W., Haarmann, A., Schaefer, F.: Combining skin conductance and heart rate variability for adaptive automation during simulated IFR flight. In: Harris, D. (ed.) HCII 2007 and EPCE 2007. LNCS (LNAI), vol. 4562, pp. 639–647. Springer, Heidelberg (2007)