

The Role of Instant Messaging on Task Performance and Level of Arousal

Sylvain Bruni

16.422 Human Supervisory Control

MAS.630 Affective Computing

Introduction

- Common project:
 - 16.422 Human Supervisory Control
 - Influence of IM on task performance
 - MAS.630 Affective Computing
 - Influence of IM on level of arousal
 - AC = “How computational systems can sense, recognize and understand human emotions and respond”

Motivation

- Instant Messaging (IM) is a collaborative communication tool
- IM enables an informal way to communicate
- IM builds bridges to other media
- BUT:
 - IM is a pervasive tool
 - IM does not provide info on people's situation
 - IM is intrusive and disruptive

Motivation

Tactical Tomahawk Dual Screen Human Supervisory Control Interface

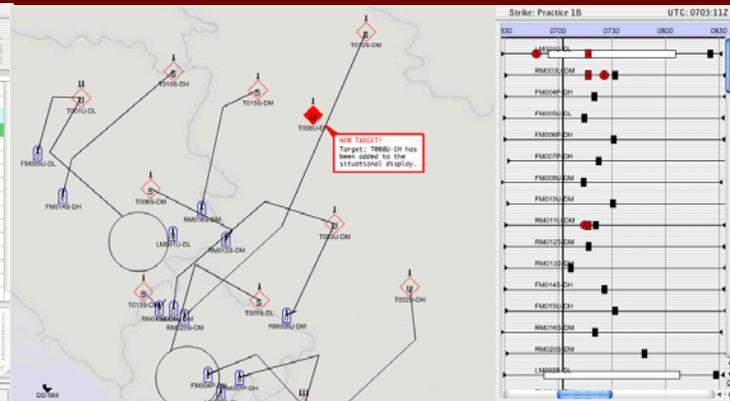
Strike: BB UTC: 0704:40Z

	Convoy	ZSU-23	Munitions Depot	Truck Park	Barracks	Electric Grid	Training Camp	SAM Site	Oil Storage	Fuel Depot
	0700	0717	0711	0714	0721	0725	0729	0731	0733	0734
	0701	0720	0702	0706	0708	0712	0716	0719	0722	0724
LM029U-DL										
RM0205-DM	5:08	---	---	---	---	16:27	16:27	---	10:32	14:41
LM0315-DL	3:47	Retargeting		29:43	24:33	13:22	---	---	---	---
RM022U-DM				0:07:25Z	3:08F	9:34	---	3:17	---	0:07:21Z
RM0315-DL	---	---		0:07:26Z	5:20	---	---	---	---	---
RM0345-DM	---	---	15:00					---	8:10	18:40
RM035U-DM								10:48	12:26	17:33
RM036U-DL										
FF Alloc	0	0	2	0	0	1	1	1	1	0
RTG Alloc	0	0	0	1	1	0	0	0	0	1
Total Required	1	1	2	1	1	1	1	1	1	1
Shortage	>> 3:40	>> 1:44	---	---	---	---	---	---	---	---

Target: T0515-EH
ZSU-23
Lat: 34.70N
Long: 52.50E
Elevation: 278ft
Missile(s) targeting: None

Misc: LM0315-DL
Type: Block IV - GPS Only
From: USS Stump (DD-978)
Launch: 06:30Z
Mission # 000015
Current Target: T0405-DL
TTL: 87m6s (08:30Z)
TOIR: 87 min
Loiter: 71 min total loiter
Net flex capable

Retarget: Retarget Missile
Retargeting command sent



Tactical Tomahawk IM Interface

Show: Action Messages Information Messages Health/Status Warnings

(07022) Blue Ridge => EMERGENT TARGET IN THE SOUTH CENTRAL SECTOR
(07022) Blue Ridge => How many missiles will hit their targets after 0730??
(07022) TTMR => Retargeting LM0315-DL to T0515-EH
(07032) Blue Ridge => How much loiter time is remaining for LM029U?
(07032) Blue Ridge => Intelligence reports a possible emerging target in the central sector.
(07042) Blue Ridge => EMERGENT TARGET IN NORTH CENTRAL SECTOR
(07042) Stump => Missile FM052P-DH impacted target
Last message received @ 0704Z

Clear Send

Motivation

- IM = series of discrete events
- Repeated discrete interruptions modify alertness / level of arousal
- Level of arousal = “how awake I am in response to an emotional stimulus”
- Alertness = “how much I am prone to give a quick response”

Experimental Design

- 6 subjects
- Play ATC sim game (primary task)
- Respond to IM (secondary task)
- Protocol:
 - Basic rules
 - Live demo
 - Questions
 - Experiment



Experimental Design



- Experiment configuration
- Skin Conductivity Response (SCR) measurement device: GSR (Galvanic Skin Response)



Skin conductivity was recorded using a galvanic skin response (GSR) measurement device disposed on the subject's left hand (if right-handed, right hand if left-handed), which was to remain motion less over the entire experiment. Here is a picture of the GSR .The skin conductivity signal was sent to the experimenter's computer using Bluetooth technology, and Python interpretation code.

Experimental Design

- 3 independent variables:
 - Gender (M/F)
 - Workload (low/high) = #planes (4/12)
 - Flow of IM (none, low, high) = relevant questions
- 3 dependent variables: Score, Time delay, SCR
- 6 scenarii

	Low WL	High WL
No IM	1	4
Low IM	2	5
High IM	3	6

Experimental Design



5/8/2004	2:36:33 PM	Sylvain	Subject	start now!
5/8/2004	2:36:48 PM	Sylvain	Subject	What time is it?
5/8/2004	2:36:56 PM	Subject	Sylvain	00:45
5/8/2004	2:37:01 PM	Sylvain	Subject	What flight is going to CCC?
5/8/2004	2:37:15 PM	Sylvain	Subject	What flight is taking off from JFK?
5/8/2004	2:37:16 PM	Subject	Sylvain	tw362h
5/8/2004	2:37:24 PM	Subject	Sylvain	qa916h
5/8/2004	2:37:33 PM	Sylvain	Subject	What is TW362H's heading?
5/8/2004	2:37:44 PM	Subject	Sylvain	ccc
5/8/2004	2:37:49 PM	Sylvain	Subject	What many planes are you responsible for?
5/8/2004	2:37:54 PM	Subject	Sylvain	4
5/8/2004	2:38:06 PM	Sylvain	Subject	What flight is going to HVN?
5/8/2004	2:38:19 PM	Subject	Sylvain	no one
5/8/2004	2:38:35 PM	Sylvain	Subject	What is flight QA916H's altitude?
5/8/2004	2:38:47 PM	Subject	Sylvain	5k
5/8/2004	2:39:06 PM	Sylvain	Subject	How many VOR are there?
5/8/2004	2:39:12 PM	Subject	Sylvain	3
5/8/2004	2:39:19 PM	Sylvain	Subject	You sure?
5/8/2004	2:39:28 PM	Subject	Sylvain	yes
5/8/2004	2:40:14 PM	Sylvain	Subject	stop now!
5/8/2004	2:41:06 PM	Sylvain	Subject	start now!
5/8/2004	2:42:45 PM	Sylvain	Subject	collision
5/8/2004	2:42:48 PM	Sylvain	Subject	stop now
5/8/2004	2:43:23 PM	Sylvain	Subject	start now!
5/8/2004	2:43:23 PM	Sylvain	Subject	stop now!
5/8/2004	2:43:23 PM	Sylvain	Subject	start now!
5/8/2004	2:43:23 PM	Sylvain	Subject	Where is going CH69?
5/8/2004	2:56:27 PM	Sylvain	Subject	msv
5/8/2004	2:56:43 PM	Subject	Sylvain	Where is going NW285?
5/8/2004	2:57:11 PM	Sylvain	Subject	jk
5/8/2004	2:57:29 PM	Subject	Sylvain	What is flight US47's altitude?
5/8/2004	2:57:54 PM	Sylvain	Subject	3k
5/8/2004	2:58:16 PM	Subject	Sylvain	What is CS37's heading?
5/8/2004	2:58:49 PM	Sylvain	Subject	ewr
5/8/2004	2:58:56 PM	Subject	Sylvain	no
5/8/2004	2:59:01 PM	Subject	Sylvain	How many planes will land at JFK?
5/8/2004	2:59:34 PM	Sylvain	Subject	3
5/8/2004	2:59:44 PM	Subject	Sylvain	Where is going GF23?
5/8/2004	3:01:04 PM	Sylvain	Subject	jk
5/8/2004	3:01:16 PM	Subject	Sylvain	How many planes are going to HVN?
5/8/2004	3:01:28 PM	Sylvain	Subject	2
5/8/2004	3:01:44 PM	Subject	Sylvain	Where is flight TW474?
5/8/2004	3:02:50 PM	Sylvain	Subject	bdr
5/8/2004	3:03:02 PM	Subject	Sylvain	How many plane still have to exit via a VOR?
5/8/2004	3:03:55 PM	Sylvain	Subject	4
5/8/2004	3:04:16 PM	Subject	Sylvain	stop now!
5/8/2004	3:04:50 PM	Sylvain	Subject	

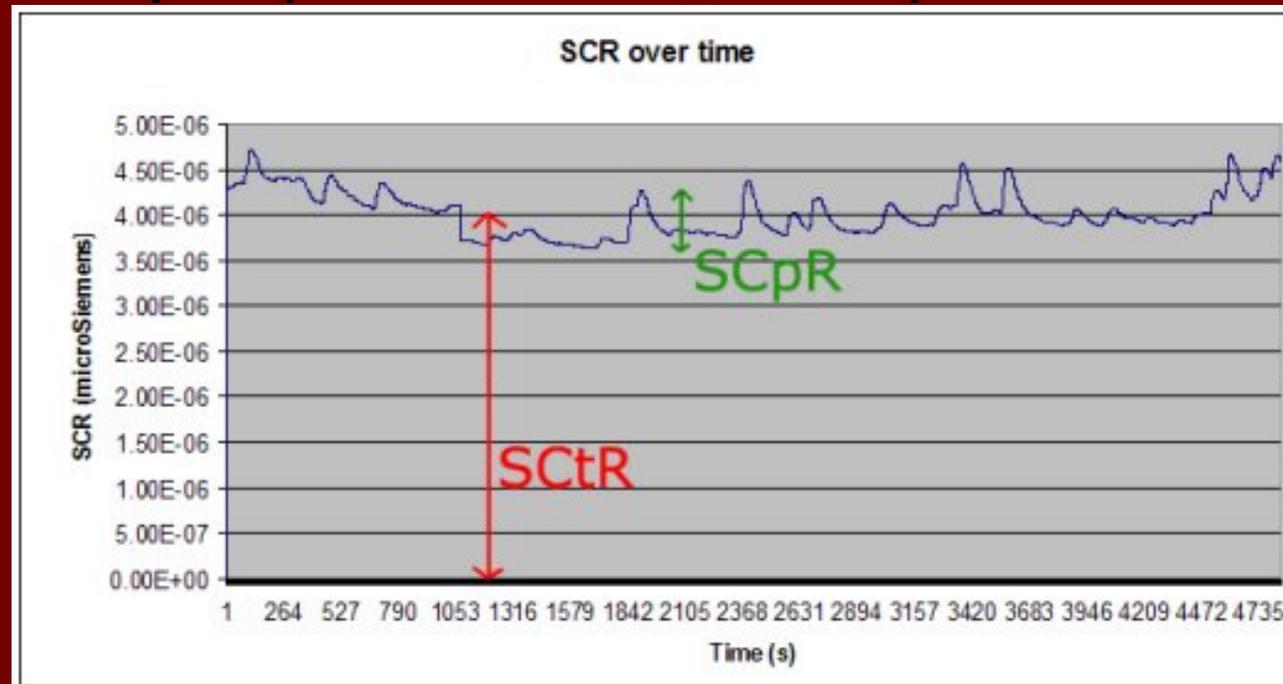
Workload was controlled by the number of planes (4 planes for the low workload case, and 12 planes for the high workload case). Following are successive caption of scenarios with 4 and then 12 planes.

Flow of instant messages was subjectively controlled by the experimenter: no IM in the "no IM" case, an IM every minute or two in the "low IM" case and a constant flow of IM during the "high IM" situation. Each message consisted in a question relative to the current situation in the game. Here is a sample of questions asked:

Results

- $\text{Score}/(\text{max score}) = \text{performance on scenario}$
- $\text{Delay} = \text{average delay over scenario}$
- $\text{Skin Conductivity} = \text{phasic} + \text{tonic components}$

- **SCtR:**
 - Tonic
 - Baseline
- **SCpR:**
 - Phasic
 - Changes



The ratio of the first two gave the task performance.

Time delays were averaged to give an average time response for each scenario. Even if the questions were of variable difficulty, the proportion and occurrence of easy and more difficult were conserved during the different scenario. This measure can therefore be interpreted as a global amount of time allocated to the task of responding to the IM. Maybe a better way to quantify time response would be to average time delays over the 4 or 5 most difficult questions.

From the skin conductivity response, two values were quantified for each scenario:

- SCtR: skin conductivity tonic response (which corresponds to the overall, global level of conductivity, typically from 0 to 10 microSiemens);

- SCpR: skin conductivity phasic response (which corresponds to the fast varying responses to particular events, ranging from 0 to 0.1 microSiemens).

In this experiment, SCpR was averaged among all the particular distinctive responses. In the scenarii with no IM, SCpR corresponded to the influence of the game and its particular events; whereas it corresponded to the impact of the game and of the incoming IM in the scenarii with IM. SCR is shown here in microSiemens.

Nevertheless, in order to perform the statistical analysis, the direct output of the GSR measurement device was used. It is given by the linear relation:

$\text{GSR_output} = 6.55 \times 10^8 \times \text{Skin_Conductance}$.

Analysis / Discussion

■ *SCtR*

- Data not normally distributed: 1 subject with abnormally high skin conductivity (> 16 microSiemens). Subject removed.
- Multiple ANOVA: only IM flow ($p < 0.014$, with 0.727 of power) significant influence on the tonic response:
 - IM is VERY intrusive: it modifies the component associated to mood and overall emotional state
- Correlation tests: *SCtR* positively correlated with *SCpR* ($p < 0.01$)
 - The higher the tonic component is, the bigger the phasic modifications will be. In this situation, people with high tonic response will have a tendency to respond even more to interruptions

■ *SCpR*

- Multiple ANOVA: IM flow ($p < 0.005$, with 0.883 of power) and workload ($p < 0.003$, with 0.908 of power) affect the phasic response.
- Post-hoc analysis: difference between no IM and high IM is extremely significant ($p < 0.004$).
- Correlation tests:
 - # *SCpR* inversely correlated with score ($p < 0.002$)
 - Subjects performing well showed less skin conductivity variation.
 - # *SCpR* positively correlated with time delay ($p < 0.022$)
 - subject delaying responses to IM had higher phasic components (stress, anxiety...)

Analysis / Discussion

■ *Delay*

- Multiple ANOVA: no statistical result appeared.
- Correlation tests: delay inversely correlated to score ($p < 0.012$).
 - Subjects performing well on the game (high scores) have more time to answer the IM, and thus have shorter delays.

■ *Score*

- Data not normally distributed: almost all subjects at 100% for easiest scenario (data skewed to the right). Scenario 1 removed.
- Multiple ANOVA:
 - workload is a significant factor ($p < 0.005$, power of 0.869)
 - IM is a significant factor ($p < 0.028$, power of 0.620)
 - gender*IM is also significant ($p < 0.037$, power = 0.568)...
 - ...but gender itself is not significant ($p = 0.581$).

Discussion

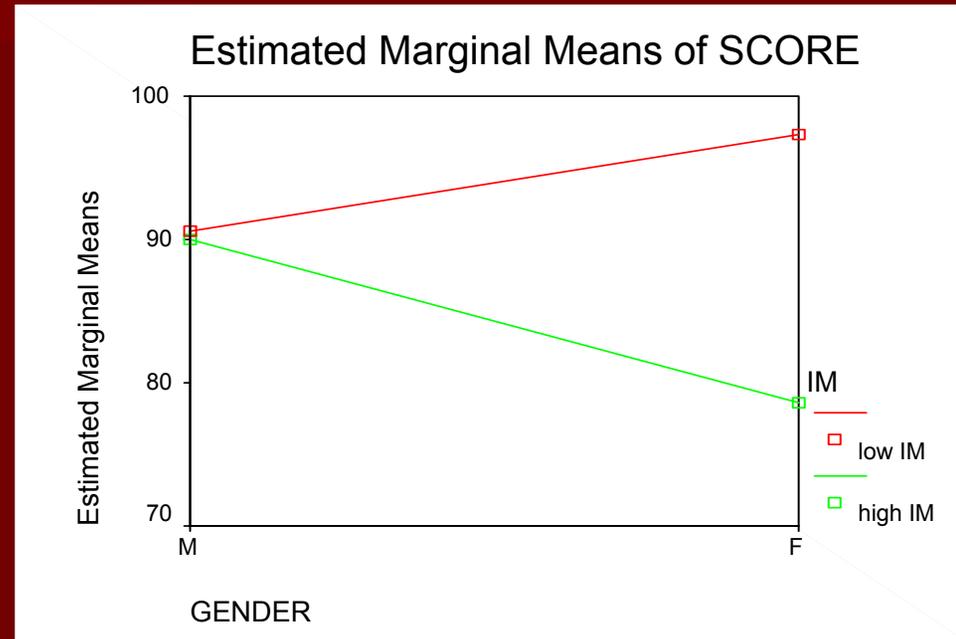
- IM influences tonic response
 - IM and WL influence phasic response
 - Tonic and phasic responses are positively correlated
 - Phasic response inversely correlated to score
 - Phasic response positively correlated to delay
- >> IM is a stressing factor (despite being informal)
- >> IM modifies mood under high pressure

Discussion

- Expected results:
 - delay inversely correlated with score
 - workload influences (badly) score
 - New results:
 - IM influences (badly) score
 - Especially significant with women
- >> IM affects significantly task performance under high workload scenarii
- >> Men and women do not handle the situation the same way

Discussion

- But be careful: only 6 subjects (4M, 2F)
- Women: the only to send “irrelevant” messages (smileys, “ok” confirmation)
- Men tend to focus more...
- ... but overall result is comparable (women perform better in low IM)



Conclusion

- Issue still needs work: more subject, more control of IM disruptiveness (level of difficulty of questions, rhythm...)
- Even if IM not answered, still influence behavior (alert sound, expectations, stress of not answering...) and task performance
- Better design of IM needed: SA indicator would be recommended

Resources

- <http://courses.media.mit.edu/2004spring/mas630/04.projects/sbruni/index.html>
- Literature review
- Presentation
- Final Paper
- GSR Socket (Python) file

- Questions?