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

Sylvain Bruni, *Massachusetts Institute of Technology, May 2004.*

***Abstract***. Despite being an informal, collaborative way to communicate, Instant Messaging (IM) remains a pervasive tool. Repeated discrete messages can be unwelcome, and the current technology does not allow for situation awareness information about the recipient of the messages. Therefore, concerns are that IM might be disruptive enough to impair performance on concurrent tasks. In this experiment, subjects were exposed to six different scenarii of an Air Traffic Control simulation game while answering instant messages. Each scenario combined two levels of workload (low and high) and three levels of flow of IM (none, low, high). Performance and time to respond were recorded, as well as skin conductivity, physiological parameter linked with level of arousal. Workload and flow of IM where shown to reduce performance. Whereas gender does not have a global influence on score, women are less robust to IM interruptions, especially under high pressure. As expected, time delay was inversely correlated to score. It also turned out that IM modified the skin conductivity response component associated with mood and overall emotional state. IM and workload have a significant impact on subjects' anxiety, which is also correlated positively with time delay and inversely to score. These results call for a better design of chat interface and a better management of instant messaging.

## *Introduction*

### *Instant Messaging*

More than 60 million users have registered to the AOL Instant Messaging (IM) services (Andrews 2001) and the increasing number of available IM clients makes this communication technology one of the most popular and expanding. Companies use IM professionally, for communication between teams located in different sites (Tang, Yankelovich and Begole 2000). The military also use IM as a collaborative, real-time, informal communication tool (Cummings 2003). Nagel (2002) refers to IM as a computer mediated conversation that allows distributed collaboration. All these ways to use and definitions are consistent with very recent research that proved that informal communication facilitates collaborative work tasks (Huang, Russel and Sue 2004). This is in part explained by the IM being flexible and expressive (Nardi, Whittaker and Bradner 2000).

But IM is pervasive, in the sense that its operational structure can lead to unwelcome distractions: the immediacy of contact permitted by IM can go against the recipient's need for privacy (Deckmyn 1999). Therefore, automatically popping-up windows are both visually disruptive and socially aggressive. Indeed, the lack of a transition or an introduction phase for a proper, polite initiation of conversation prevents IM from following standard social rules. In other words, IM is disruptive because of the absence of direct verbal feedback, as in real conversations (Tang, Yankelovich and Begole 2000).

IM gives the users "a sense of ultra-compressed time, and foreshortened horizons" (McKenna 1997). Some IM users integrate this as an additional time constraint, that increases time pressure: they feel required to adapt to this high-pace time pattern generated by IM. Past experiments have proved that primary tasks performed under consequent time-pressure, are significantly degraded (Cellier and Eyrolle 1992).

*Performance*

Cummings and Guerlain (2003) confronted an interface problem during their experiments: subjects tended to focus on the IM chat box interface, while leaving their primary task, which led to their loss of situation awareness and a degradation of performance. Moreover, in the case of ATC and pilot communication, such interruptions have been shown to affect considerably performance on both sides, and thus to affect the overall safety (Latorella 1998).

Several past studies have also showed that vigilance level in human supervisory control tasks dropped deeply during the first thirty minutes of watch in a phenomenon referred to as the vigilance decrement (Mackworth 1948; Harris and Chaney 1969; Parasuraman 1986). Therefore, it is legitimate to think that IM, simultaneously with the unavoidable vigilance decrement, can have a tendency to worsen it, by interfering with ongoing cognitive workload. Workload refers to the "cost of accomplishing task requirements for a human involved in a man-machine system" (Hart and Wickens 1990). With this in mind, IM can be thought as interruption processes that sneak into the current workload. Therefore, interruption management should be of primary concern in traditional task management processes. Latorella (1998) defines interruption management as the ability to "attend appropriately to and to accommodate new, interrupting stimuli and tasks".

Interruptions that occur after important tasks or between non-dividable subtasks are less harmful (Czerwinski, Cutrell and Horvitz 2000). This is consistent with the "chunking behavior" initially introduced by Miyata and Norman (1986): tasks consist in a succession of subtasks or chunks that cannot be individually interrupted. The user first finishes the current task chunk before switching to the interruption.

Therefore, with the increasing use of instant messaging as a communication tool, concerns have raised that it was disruptive enough to degrade one's performance on concurrent tasks.

*Level of Arousal*

IM can be considered as a series of discrete events. Therefore, it can be expected that these repeated discrete interruptions modify alertness and/or level of arousal. Level of arousal is understood here as “how awake [the subject is] in response to an emotional stimulus”, and alertness as “how much [the subject is] prone to give a quick response”. Level of arousal is usually measured through skin conductivity.

The skin conductivity response consists of two components: the tonic and phasic (Boucsein 1992). The tonic component is slow moving, oscillating over the course of days, whereas the phasic component is fast moving, and spikes sharply when a person is startled, and generally increases when a person is psychologically aroused.

In other words, the tonic component of the skin conductivity response corresponds to the overall mood, whereas the phasic component corresponds to the anxiety or stress felt in result to a particular situation.

In the following experiment, it is hypothesized that instant messaging will decrease overall performance (decrease score, and increase time delays), as well as increase the phasic component of the skin conductivity response. It is expected that workload has the same influence.

### *Methods and Experimental Design*

The goal of the present experiment is to observe the impact of three independent variables (gender, workload, flow of IM) on three dependent variables (performance, time delays, skin conductivity response).

6 subjects were involved in this experiment. They underwent a series of 6 different scenarii consisting in playing an Air Traffic Control (ATC) simulation game (Air Command 3.0, Shrapnel Games) while responding to incoming instant messages (through an MSN Messenger 6.1 chat interface). The instruction was made clear: the subjects were required to play the game and consider it as their primary task, and respond to instant messages when they could (secondary task).

They were first explained basic knowledge on how to play the game as well as the rules to follow, and then shown a demo to see a live example of a simulation scenario. Figure 1 shows a caption of the ATC simulation game interface.

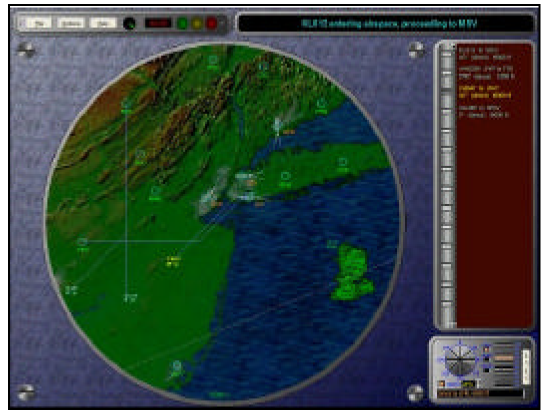


Figure 1. Caption of the ATC simulation game interface (Air Command 3.0, Shrapnel Games).

The game was available on an independent laptop computer. A second computer was used to provide the chat interface (Figure 2).

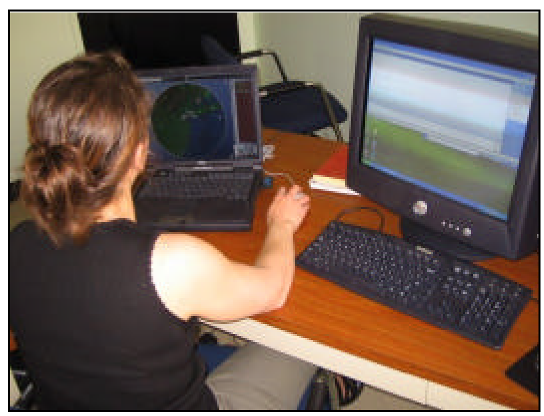


Figure 2. Subject during an experiment. Left computer is dedicated to the ATC simulation game, right computer to instant messaging.

During the experiment, score at the game was recorded, as well as maximum possible score. In addition, time delays for responses to the instant messages was made available by the historical listings of MSN Messenger 6.1. In order to measure level of arousal, 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 motionless over the entire experiment. Figure 3 shows the GSR device. Two electrodes placed at precise locations on the hand allowed for the measurement of skin resistance. The electrodes were linked to a small electronic circuit, transforming skin resistance into skin conductivity.

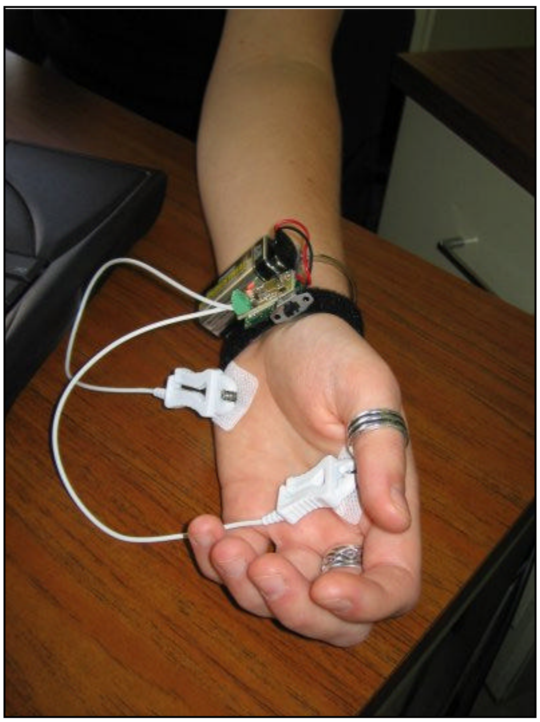


Figure 3. The Galvanic Skin Response measurement device connected on a subject's left hand.

The skin conductivity signal was sent to the experimenter's computer using Bluetooth technology, and a Python interpretation code.

The 6 scenarii played combined two levels of workload (low and high) and three levels of flow of IM (none, low, high). Workload was controlled by the number of planes (4 planes for the low workload case, and 12 planes for the high workload case). Figure 4 shows a caption of a scenario with 4 planes, figure 5 shows a caption of a scenario with 12 planes.

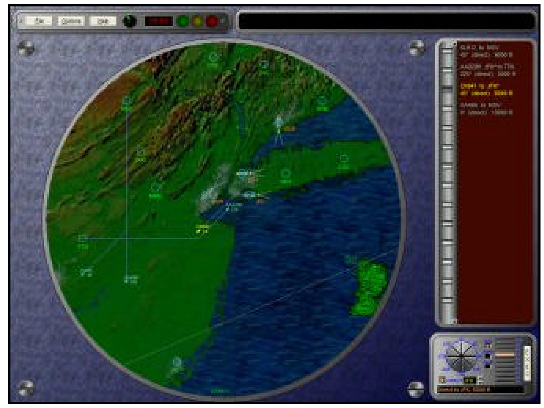


Figure 4. Caption of a scenario with 4 planes.

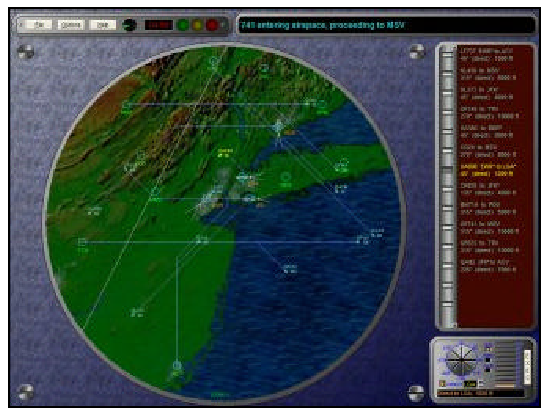


Figure 5. Caption of a scenario with 12 planes.

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, such as: "How many planes will land at JFK?"; "What is flight AA952's altitude?"; "Where is going flight IB34H?". When an IM was received, the subject could hear a characteristic tone.

The order of the scenarii was the same for all subjects:

1- low WL - no IM
2- low WL - low IM
3- low WL - high IM
4- high WL - no IM
5- high WL - low IM
6- high WL - high IM

This specific order was determined by a pilot study: since the game automatically shuts down when a collision occurs, it was preferable to avoid collisions as much as possible, therefore increasing the difficulty progressively, which is the case with this protocol.

*Measurements*

For each of the 6 scenarii, 4 measures were taken:
- score on the scenario;
- maximum score possible for the scenario;
- time delays to respond to the IM (except in the scenario with no IM);
- skin conductivity.

The ratio of the first two gave the task performance, in percentage.

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 questions were conserved during the different scenarii. This measure can therefore be interpreted as a global amount of time allocated to the task of responding to the IM.

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 variating 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. Figure 6 presents a typical skin conductivity response.

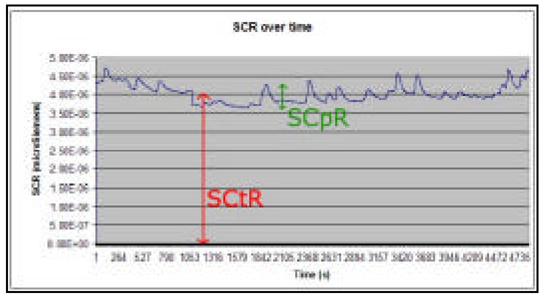



Figure 6. Typical skin conductivity response (SCR, in blue) with its two components: the tonic (SCtR, in red) and the phasic (SCpR, in green) components. Units are time in second and SCR in microSiemens.

*Note*: in order to perform the statistical analysis, the direct output of the GSR measurement device was used. It is given by the linear relation:
GSRoutput = $6.55 \times 10^8$ x SCR.

***Results.***

*SCtR*

A one sample Kolmogorov-Smirnov test showed that the data was not normally distributed, because of a subject that had abnormally very high skin conductivity (more than 16 microSiemens). Therefore, this subject was removed for the analysis of skin conductivity.

A multiple ANOVA was performed to

find out the influence of gender, workload level and IM flow on SCtR. Only one independent parameter showed to affect SCtR significantly: IM flow ($p<0.014$, with 0.727 of power).

A set of correlation tests (Pearson correlation, Kendall's tau and Spearman's rho non parametric correlations) was performed. A very significant result ($p<0.01$) appeared: SCtR had a tendency to increase with SCpR, meaning that, the higher the tonic component is, the bigger the phasic modifications will be.

*SCpR*

A multiple ANOVA was performed to find out the influence of gender, workload level and IM flow on SCpR. Two independent parameters showed to affect SCtR significantly: IM flow ($p<0.005$, with 0.883 of power) and workload ($p<0.003$, with 0.908 of power). Post-hoc analysis showed that the difference between no IM and high IM is extremely significant ($p<0.004$).

A set of correlation tests (Pearson correlation, Kendall's tau and Spearman's rho non parametric correlations) was performed. The first very significant result ($p<0.002$) that appeared was with score: SCpR had a tendency to decrease when score increased, meaning that, subjects performing well showed less skin conductivity variations. The second significant result was with delay ($p<0.022$). SCpR has a tendency to increase with time delay: when a subject delayed its responses to IM, its phasic component had a tendency to be higher.

*Delay*

A multiple ANOVA test was run, but no statistical result appeared. A set of correlation tests was performed (Pearson correlation, Kendall's tau and Spearman's rho non parametric correlations). It turned out that delay was inversely correlated to score ($p<0.012$). This was expected: subjects performing well on the game (high scores) would have more time to answer the IM, and thus have shorter delays.

*Score*

A one sample Kolmogorov-Smirnov test showed that the data was not normally distributed, because almost all subjects had a score of 100% for the easiest scenario. This made the data really skewed to the right. Therefore, the data for scenario 1 (with low workload and no IM) was removed. The remaining data was normal enough to perform the following tests.

A multiple ANOVA test was performed. Three results came out positive: workload was a significant factor ($p<0.005$ with power at 0.869); IM was significant ($p<0.028$ and power of 0.620) and gender*IM was also significant ($p<0.037$, power = 0.568). This last result was unexpected, especially since gender itself is not significant ($p=0.581$).

***Discussion***

These results have to be considered very carefully. Only 6 subjects participated so far in this experiment, preventing therefore from drawing infallible conclusions. Nevertheless, some trends and considerations can be observed.

The skin conductivity data has showed that: (1) IM flow influences the tonic component of the skin conductivity response, (2) IM flow and workload influence the phasic component of the skin conductivity response, (3) tonic and phasic responses are positively correlated, (4) the phasic response is inversely correlated to score, and (5) the phasic response is positively correlated to delay.

Conclusion (1) implies that IM has an influence on the mood of the subject, which usually varies on durations of the order of days to months. Here, mood variation occurred on the order of minutes, which confirms the fact that IM is emotionally pervasive, especially in high pressure scenarii.

Conclusion (2) shows that both IM and workload are stressing factors: this was expected for workload (the more the subject has to do, the more stress he/she will be). IM induces anxiety in the user for several reasons: being disrupted during an important task elicits natural anxiety, and knowing that instant messages have been received, but are left unanswered, contributes to making it bigger. Even if the primary task is attended correctly, the influence of IM is real.

Conclusion (3) basically states that people with high skin conductivity at rest will experience bigger physiological reactions in response to particular events.

Conclusion (4) and (5) state expected results: the better a subject performed (high score, and short delays), the lower he/she physiologically reacted (phasic component). The subject being in a peaceful state of mind was retranscribed in his/her physiological response. The other way around allows the possibility for detecting when an operator is in a difficult situation: the phasic component can become a characteristic alerting signal.

Overall, skin conductivity tells that IM is a stressing factor, despite it's being informal. Figures 7 and 8 show respectively the SCR (GSR output) for the same subject in scenarii 2 and 3. It is quite obvious that the skin conductivity response is much less smooth in scenario 3, when the flow of IM is very high. This shows how much the subject is dedicated to the task, and puts resources to it. The subject is then much more prone to react, his/her level of arousal is globally higher.

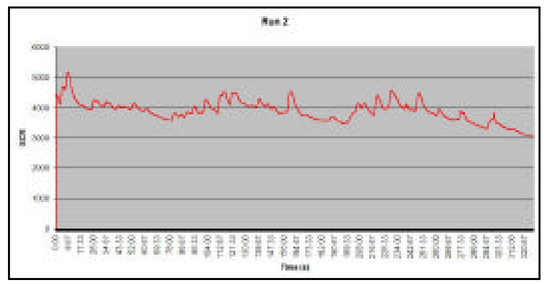


Figure 7. GSR output for scenario 2 (low workload, low IM).

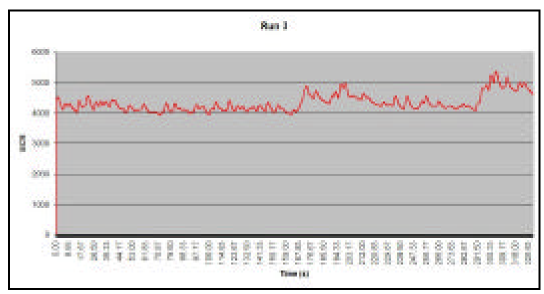


Figure 8. GSR output for scenario 3 (low workload, high IM).

Alerting sounds, expectations, stress of not answering may be potential factors modifying the level of arousal of the operator, and distracting stimuli affecting attendance to the primary task.

Concerning delays and scores, expected results showed up. Delay is inversely correlated with score, which is consistent with the experimental set up: the better the subject plays to the game, the shorter the delays will be to respond to the messages. The negative influence of workload on score was also noticed: the more difficult the situation is, the worse the score will be. The important results are that IM influences score (the bigger the flow of IM is, the worse the score gets) and this is especially significant with women under high flows of IM. Even though only two women participated in the experiment, it is worth noticing that they were the only subjects to send "irrelevant" messages (as smileys or unnecessary "ok" of confirmation). Men tended to focus better on the primary task, hence the similar results in the two IM situations, whereas women were particularly affected by the high flow of IM (figure 9). However, the overall performance was comparable. In addition, the very small sample tested is not enough to draw strong conclusions about the gender*IM effect.

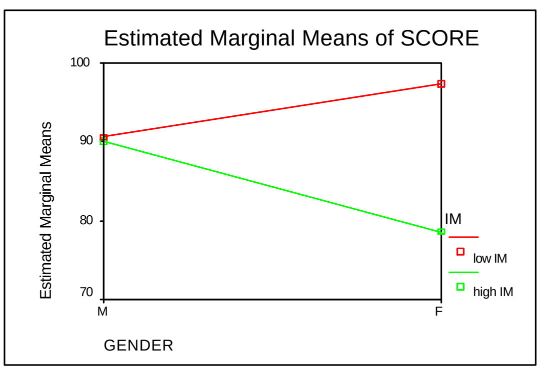



Figure 9. Mean score vs. gender, for the two IM situations (low flow of IM, red; high flow of IM, green).

## *Conclusion*

Instant messaging is a wide-spread communication tool, but its impact on task performance is not well known. This experiment uncovered some trends: even if not answered, instant messages influence the operator's behavior and task performance.

Indeed, IM reduces task performance, especially under high workload scenarii. However, men and women do not handle the situation the same way: men are more robust to distractions in such a context. In addition, it turned out that IM is a stressing factor whose effect is to modify both the overall level of arousal and the discrete emotional responses to particular events; but further investigation is needed to see if men are, in general, more robust to modifications of level of arousal.

If IM is to be kept as a communication tool in high pressure environments, a better design of its interface may be needed: for example, a situation awareness indicator could be implemented. Its role would be to inform the IM sender about the level of workload or stress the other is undergoing. Following the experiment described here, such an indicator could be based on the phasic response of the operator's skin conductivity.

This issue still needs work: more subject have to be tested in order to draw solid conclusions. Especially, it would be interesting to investigate more the gender*IM effect, to see if it really is robust on a large scale. This experiment could also be improved with more control of IM disruptiveness: level of difficulty of questions, rhythm of flow might be subfactors interesting to look at.

## *References*


**Andrews P** (2001). Instant Messaging: Lessons in Pervasive Computing. IBM Executive Tek Report, April 16, 2001.

**Boucsein W** (1992) Electrodermal Activity, Plenum Series in Behavioral Psychophysiology and Medicine, Plenum Press.

**Cellier J and Eyrolle H** (1992). Interference between switched tasks. Ergonomics, 35, 25-36.

**Cummings ML** (2003). Tactical Tomahawk Interface for Monitoring and Retargeting. Presented at the Precision Strike Conference, JHU/APL, October 15, 2003.

**Cummings ML, Guerlain S** (2003). Evaluating Situation Awareness through an Embedded Instant Messaging Interface.

**Czerwinski M, Cutrell EB and Horvitz E** (2000). Instant Messaging: Effect of Relevance and Time. In S. Turner and P. Turner (Eds). People and Computer XIV: Proceedings of HCI 2000, Vol. 2, British Computer Society, pp. 71-76.

**Deckmyn D** (1999). Valuable Tool or Distraction? IT debates merits of instant messaging. Computer World, Vol. 33, No 28, p.1.

**Harris DH, Chaney FD** (1969). Human factors in quality assurance. New-York: Willey. (Reprinted 1988, Los Altos, CA: Peninsula).

**Hart SG, Wickens CD** (1990). Workload Assessment and Prediction. In HR Booher (Ed). Manprint: An integrated approach to systems integration (pp. 257-296). New-York: Van Nostrand.

**Huang EM, Russell DM and Sue AE** (2004). IM Here: Public Instant Messaging on Large, Shared Displays for Workgroup Interactions. CHI 2004. Vienna, Austria.

**Latorella KA** (1998). Effects of Modality on Interrupted Flight Deck Performance: Implication for Data Link. Nasa Langley Research Center Report.

**Mackworth NH** (1948). The breakdown of vigilance during prolonged visual search. Quarterly Journal of Experimental Psychology, 1, 5-61.

**McKenna R** (1997). Real Time. Harvard Business School Publishing.

**Miyata Y and Norman DA** (1986), Psychological Issues in Support of Multiple Activities. In SW Draper (Ed.), User Centered System Design: New Perspectives on Human-Computer Interaction (pp. 265-284). Hillsdale, NJ: Lawrence Erlbaum.

**Nagel KS** (2002). Understanding Situation-Aware Interpersonal Communication. CSCW'02. New Orleans, Louisiana, USA.

**Nardi BA, Whittaker S and Bradner E** (2000). Interaction and Outeraction: Instant Messaging in Action. Proc. CSCW'00, 79-88.

**Parasuraman R** (1986). Vigilance, monitoring and search. In K. Boff, L. Kaufman and J. Thomas (Eds), Handbook of perception and human performance. Vol. 2: Cognitive processes and performance (pp. 43.1-43.39). New-York: Wiley.

**Peddemors AJH, Lankhorst MM and de Heer J** (2003). Presence, Location, and Instant Messaging in a Context-Aware Application Framework. In MS Chen et al. (Eds). MDM 2003, LNCS 2574, pp. 325-330. Springer-Verlag Berlin Heidelberg.

**Ruff HA, Narayanan S and Draper MH** (2002). Human Interaction with Levels of Automation and Decision-Aid Fidelity in the Supervisory Control of Multiple Simulated Unmanned Air Vehicles. Presences, 11(4), 335-351.

**Tang JC, Yankelovich N and Begole J** (2000). ConNexus: Instant Messaging for the Workplace: Sun Microsystems Laboratories Report.

**Wickens CD, Hollands JG** (2000). Engineering Psychology and Human Performance Third Edition, Upper Saddle River, NJ: Prentice-Hall Inc.

## *SCR Figures.*

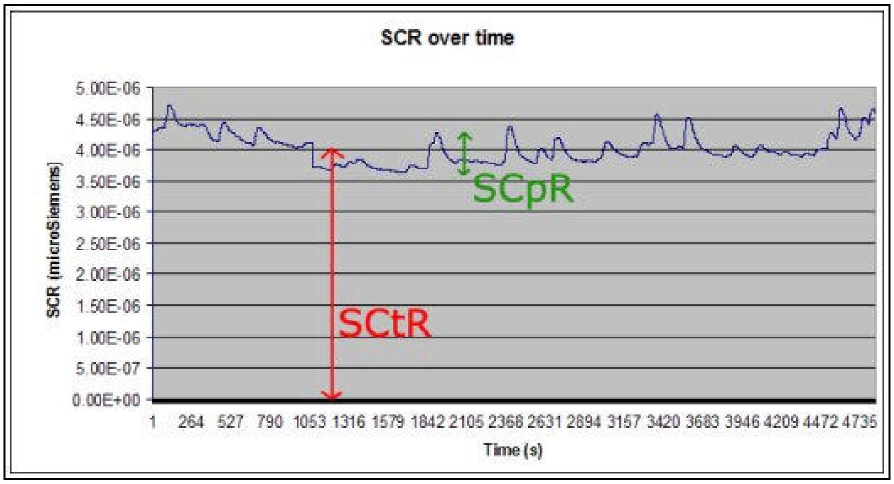



Figure 6. SCR components: tonic (SCtR) and phasic (SCpR).

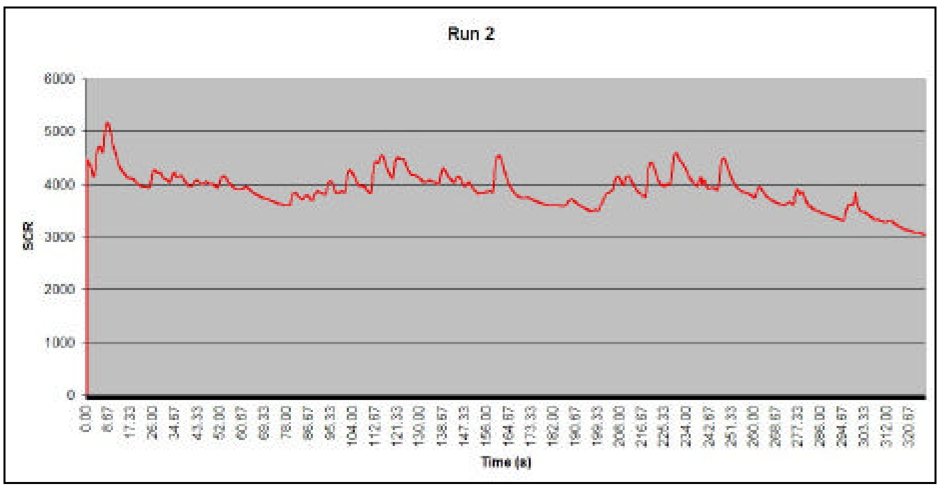



Figure 7. Scenario 2: low workload and low IM.

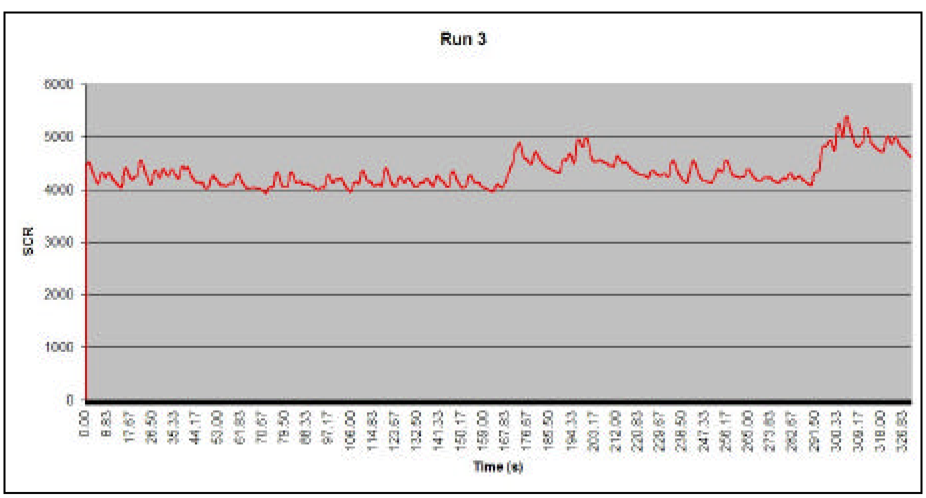
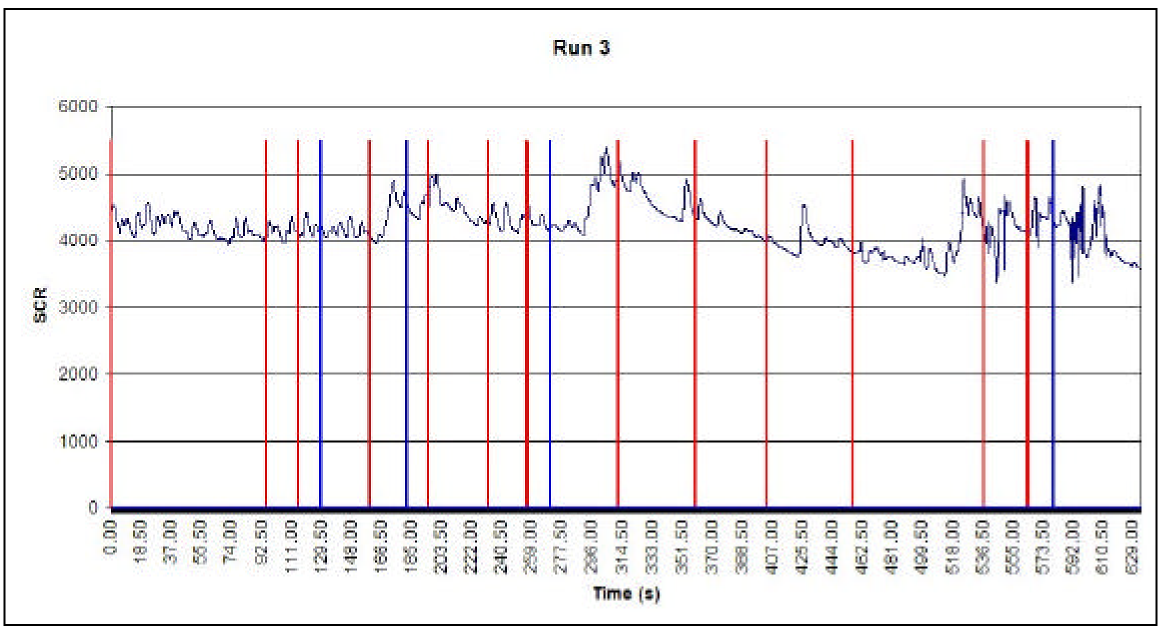



Figure 8. Scenario 3: low workload and high IM.


SCR for scenario 2 (low workload, low IM). Questions are in red, answers are in blue.




SCR for scenario 3 (low workload, high IM). Questions are in red, answers are in blue.

## *Data Analysis (main statistical outputs)*

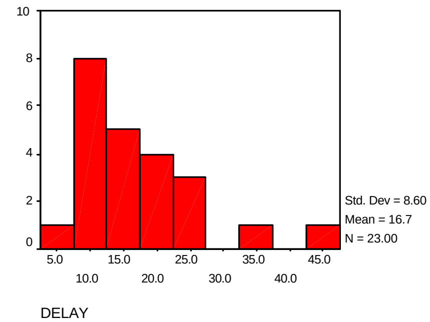
# Delay

**One-Sample Kolmogorov-Smirnov Test**

| | | DELAY |
|---|---|---|
| N | | 23 |
| Normal Parameters[a,b] | Mean | 16.6770 |
| | Std. Deviation | 8.60092 |
| Most Extreme Differences | Absolute | .186 |
| | Positive | .186 |
| | Negative | -.121 |
| Kolmogorov-Smirnov Z | | .892 |
| Asymp. Sig. (2-tailed) | | .404 |

a. Test distribution is Normal.

b. Calculated from data.




**Tests of Between-Subjects Effects**

Dependent Variable: DELAY

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Corrected Model | 576.476[a] | 7 | 82.354 | 1.175 | .372 |
| Intercept | 4966.413 | 1 | 4966.413 | 70.882 | .000 |
| GENDER | 19.647 | 1 | 19.647 | .280 | .604 |
| WORKLOAD | 62.773 | 1 | 62.773 | .896 | .359 |
| IM | 105.848 | 1 | 105.848 | 1.511 | .238 |
| GENDER * WORKLOAD | 169.720 | 1 | 169.720 | 2.422 | .140 |
| GENDER * IM | 40.477 | 1 | 40.477 | .578 | .459 |
| WORKLOAD * IM | 21.464 | 1 | 21.464 | .306 | .588 |
| GENDER * WORKLOAD * IM | 19.130 | 1 | 19.130 | .273 | .609 |
| Error | 1050.992 | 15 | 70.066 | | |
| Total | 8024.249 | 23 | | | |
| Corrected Total | 1627.468 | 22 | | | |

a. R Squared = .354 (Adjusted R Squared = .053)

**Correlations**

| | | DELAY | SCPR | SCTR | SCORE |
|---|---|---|---|---|---|
| DELAY | Pearson Correlation | 1 | .140 | -.180 | -.513* |
| | Sig. (2-tailed) | . | .524 | .410 | .012 |
| | N | 23 | 23 | 23 | 23 |
| SCPR | Pearson Correlation | .140 | 1 | .840** | -.226 |
| | Sig. (2-tailed) | .524 | . | .000 | .186 |
| | N | 23 | 36 | 36 | 36 |
| SCTR | Pearson Correlation | -.180 | .840** | 1 | .030 |
| | Sig. (2-tailed) | .410 | .000 | . | .860 |
| | N | 23 | 36 | 36 | 36 |
| SCORE | Pearson Correlation | -.513* | -.226 | .030 | 1 |
| | Sig. (2-tailed) | .012 | .186 | .860 | . |
| | N | 23 | 36 | 36 | 36 |

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

## Score

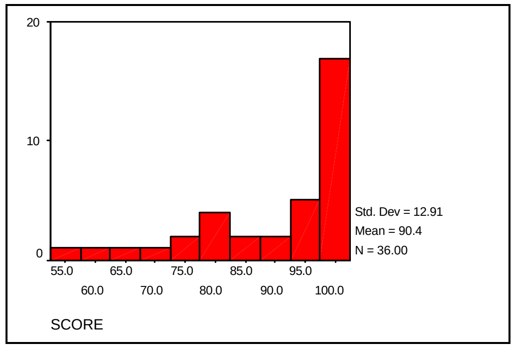
With all data:

**One-Sample Kolmogorov-Smirnov Test**

| | | SCORE |
|---|---|---|
| N | | 36 |
| Normal Parameters[a,b] | Mean | 90.3575 |
| | Std. Deviation | 12.91155 |
| Most Extreme Differences | Absolute | .259 |
| | Positive | .228 |
| | Negative | -.259 |
| Kolmogorov-Smirnov Z | | 1.552 |
| Asymp. Sig. (2-tailed) | | .016 |

a. Test distribution is Normal.

b. Calculated from data.



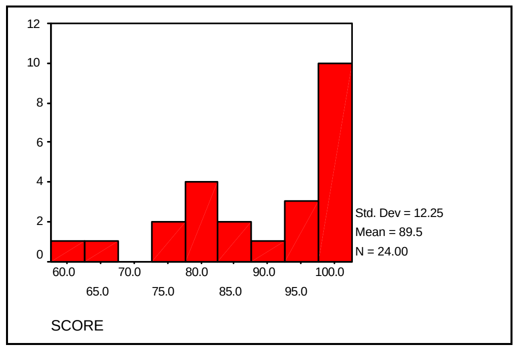

Without scenario 1:

**One-Sample Kolmogorov-Smirnov Test**

| | | SCORE |
|---|---|---|
| N | | 24 |
| Normal Parameters[a,b] | Mean | 89.4917 |
| | Std. Deviation | 12.24732 |
| Most Extreme Differences | Absolute | .256 |
| | Positive | .195 |
| | Negative | -.256 |
| Kolmogorov-Smirnov Z | | 1.255 |
| Asymp. Sig. (2-tailed) | | .086 |

a. Test distribution is Normal.

b. Calculated from data.




**Tests of Between-Subjects Effects**

Dependent Variable: SCORE

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Noncent. Parameter | Observed Power[a] |
|---|---|---|---|---|---|---|---|
| Corrected Model | 2075.002[b] | 7 | 296.429 | 3.450 | .019 | 24.147 | .850 |
| Intercept | 169415.932 | 1 | 169415.932 | 1971.494 | .000 | 1971.494 | 1.000 |
| GENDER | 27.331 | 1 | 27.331 | .318 | .581 | .318 | .083 |
| WORKLOAD | 926.115 | 1 | 926.115 | 10.777 | .005 | 10.777 | .869 |
| IM | 499.488 | 1 | 499.488 | 5.813 | .028 | 5.813 | .620 |
| GENDER * WORKLOAD | 25.521 | 1 | 25.521 | .297 | .593 | .297 | .081 |
| GENDER * IM | 442.382 | 1 | 442.382 | 5.148 | .037 | 5.148 | .568 |
| WORKLOAD * IM | 145.673 | 1 | 145.673 | 1.695 | .211 | 1.695 | .232 |
| GENDER * WORKLOAD * IM | 1.408 | 1 | 1.408 | .016 | .900 | .016 | .052 |
| Error | 1374.924 | 16 | 85.933 | | | | |
| Total | 195660.128 | 24 | | | | | |
| Corrected Total | 3449.926 | 23 | | | | | |

a. Computed using alpha = .05

b. R Squared = .601 (Adjusted R Squared = .427)

**Correlations**

| | | SCORE | DELAY | SCTR | SCPR |
|---|---|---|---|---|---|
| SCORE | Pearson Correlation | 1 | -.513* | .044 | -.215 |
| | Sig. (2-tailed) | . | .012 | .840 | .313 |
| | N | 24 | 23 | 24 | 24 |
| DELAY | Pearson Correlation | -.513* | 1 | -.180 | .140 |
| | Sig. (2-tailed) | .012 | . | .410 | .524 |
| | N | 23 | 23 | 23 | 23 |
| SCTR | Pearson Correlation | .044 | -.180 | 1 | .848** |
| | Sig. (2-tailed) | .840 | .410 | . | .000 |
| | N | 24 | 23 | 24 | 24 |
| SCPR | Pearson Correlation | -.215 | .140 | .848** | 1 |
| | Sig. (2-tailed) | .313 | .524 | .000 | . |
| | N | 24 | 23 | 24 | 24 |

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

**Correlations**

| | | | SCORE | DELAY | SCTR | SCPR |
|---|---|---|---|---|---|---|
| Kendall's tau_b | SCORE | Correlation Coefficient | 1.000 | -.368* | .051 | -.276 |
| | | Sig. (2-tailed) | . | .019 | .740 | .070 |
| | | N | 24 | 23 | 24 | 24 |
| | DELAY | Correlation Coefficient | -.368* | 1.000 | -.099 | .209 |
| | | Sig. (2-tailed) | .019 | . | .509 | .162 |
| | | N | 23 | 23 | 23 | 23 |
| | SCTR | Correlation Coefficient | .051 | -.099 | 1.000 | .507** |
| | | Sig. (2-tailed) | .740 | .509 | . | .001 |
| | | N | 24 | 23 | 24 | 24 |
| | SCPR | Correlation Coefficient | -.276 | .209 | .507** | 1.000 |
| | | Sig. (2-tailed) | .070 | .162 | .001 | . |
| | | N | 24 | 23 | 24 | 24 |
| Spearman's rho | SCORE | Correlation Coefficient | 1.000 | -.457* | .071 | -.337 |
| | | Sig. (2-tailed) | . | .028 | .740 | .108 |
| | | N | 24 | 23 | 24 | 24 |
| | DELAY | Correlation Coefficient | -.457* | 1.000 | -.181 | .306 |
| | | Sig. (2-tailed) | .028 | . | .409 | .155 |
| | | N | 23 | 23 | 23 | 23 |
| | SCTR | Correlation Coefficient | .071 | -.181 | 1.000 | .703** |
| | | Sig. (2-tailed) | .740 | .409 | . | .000 |
| | | N | 24 | 23 | 24 | 24 |
| | SCPR | Correlation Coefficient | -.337 | .306 | .703** | 1.000 |
| | | Sig. (2-tailed) | .108 | .155 | .000 | . |
| | | N | 24 | 23 | 24 | 24 |

*. Correlation is significant at the .05 level (2-tailed).

**. Correlation is significant at the .01 level (2-tailed).

**SCtR**

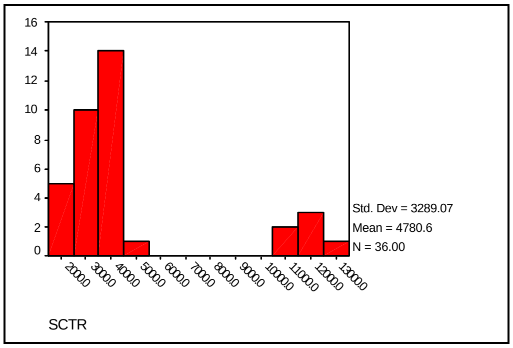
With all data:

**One-Sample Kolmogorov-Smirnov Test**

| | | SCTR |
|---|---|---|
| N | | 36 |
| Normal Parameters[a,b] | Mean | 4780.61 |
| | Std. Deviation | 3289.073 |
| Most Extreme Differences | Absolute | .364 |
| | Positive | .364 |
| | Negative | -.185 |
| Kolmogorov-Smirnov Z | | 2.184 |
| Asymp. Sig. (2-tailed) | | .000 |

a. Test distribution is Normal.

b. Calculated from data.



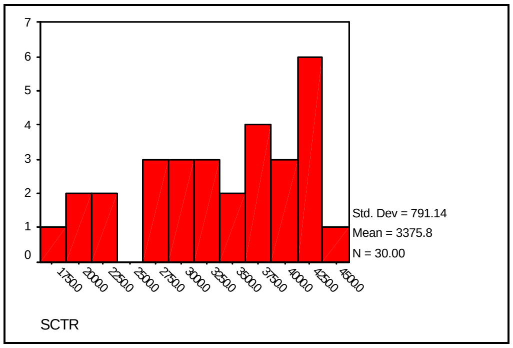

Without subject 6:

**One-Sample Kolmogorov-Smirnov Test**

| | | SCTR |
|---|---|---|
| N | | 30 |
| Normal Parameters[a,b] | Mean | 3375.83 |
| | Std. Deviation | 791.138 |
| Most Extreme Differences | Absolute | .110 |
| | Positive | .094 |
| | Negative | -.110 |
| Kolmogorov-Smirnov Z | | .600 |
| Asymp. Sig. (2-tailed) | | .864 |

a. Test distribution is Normal.

b. Calculated from data.




**Tests of Between-Subjects Effects**

Dependent Variable: SCTR

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Noncent. Parameter | Observed Power[a] |
|---|---|---|---|---|---|---|---|
| Corrected Model | 7456311.667[b] | 11 | 677846.515 | 1.141 | .388 | 12.549 | .415 |
| Intercept | 328749673 | 1 | 328749672.7 | 553.308 | .000 | 553.308 | 1.000 |
| GENDER | 5500.139 | 1 | 5500.139 | .009 | .924 | .009 | .051 |
| WORKLOAD | 4363959.606 | 1 | 4363959.606 | 7.345 | .014 | 7.345 | .727 |
| IM | 1941358.144 | 2 | 970679.072 | 1.634 | .223 | 3.267 | .299 |
| GENDER * WORKLOAD | 36266.806 | 1 | 36266.806 | .061 | .808 | .061 | .056 |
| GENDER * IM | 22351.078 | 2 | 11175.539 | .019 | .981 | .038 | .052 |
| WORKLOAD * IM | 499675.078 | 2 | 249837.539 | .420 | .663 | .841 | .108 |
| GENDER * WORKLOAD * IM | 411573.078 | 2 | 205786.539 | .346 | .712 | .693 | .097 |
| Error | 10694758.5 | 18 | 594153.250 | | | | |
| Total | 360038591 | 30 | | | | | |
| Corrected Total | 18151070.2 | 29 | | | | | |

a. Computed using alpha = .05

b. R Squared = .411 (Adjusted R Squared = .051)

**Correlations**

| | | DELAY | SCPR | SCTR | SCORE |
|---|---|---|---|---|---|
| DELAY | Pearson Correlation | 1 | .140 | -.180 | -.513* |
| | Sig. (2-tailed) | . | .524 | .410 | .012 |
| | N | 23 | 23 | 23 | 23 |
| SCPR | Pearson Correlation | .140 | 1 | .840** | -.226 |
| | Sig. (2-tailed) | .524 | . | .000 | .186 |
| | N | 23 | 36 | 36 | 36 |
| SCTR | Pearson Correlation | -.180 | .840** | 1 | .030 |
| | Sig. (2-tailed) | .410 | .000 | . | .860 |
| | N | 23 | 36 | 36 | 36 |
| SCORE | Pearson Correlation | -.513* | -.226 | .030 | 1 |
| | Sig. (2-tailed) | .012 | .186 | .860 | . |
| | N | 23 | 36 | 36 | 36 |

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

## SCpR

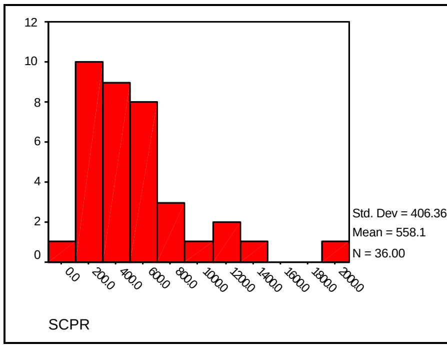
With all data:

**One-Sample Kolmogorov-Smirnov Test**

| | | SCPR |
|---|---|---|
| N | | 36 |
| Normal Parameters[a,b] | Mean | 558.06 |
| | Std. Deviation | 406.364 |
| Most Extreme Differences | Absolute | .159 |
| | Positive | .159 |
| | Negative | -.129 |
| Kolmogorov-Smirnov Z | | .957 |
| Asymp. Sig. (2-tailed) | | .319 |

a. Test distribution is Normal.

b. Calculated from data.



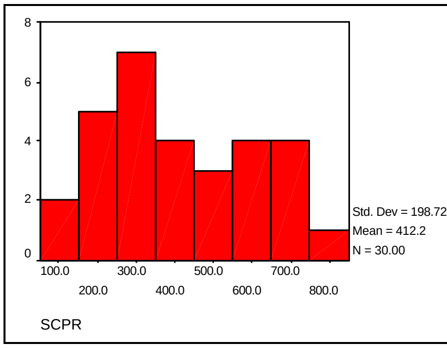

Without subject 6:

**One-Sample Kolmogorov-Smirnov Test**

| | | SCPR |
|---|---|---|
| N | | 30 |
| Normal Parameters[a,b] | Mean | 412.20 |
| | Std. Deviation | 198.715 |
| Most Extreme Differences | Absolute | .147 |
| | Positive | .147 |
| | Negative | -.126 |
| Kolmogorov-Smirnov Z | | .806 |
| Asymp. Sig. (2-tailed) | | .535 |

a. Test distribution is Normal.

b. Calculated from data.




**Tests of Between-Subjects Effects**

Dependent Variable: SCPR

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Noncent. Parameter | Observed Power[a] |
|---|---|---|---|---|---|---|---|
| Corrected Model | 772280.300[b] | 11 | 70207.300 | 3.389 | .011 | 37.282 | .926 |
| Intercept | 5028373.472 | 1 | 5028373.472 | 242.744 | .000 | 242.744 | 1.000 |
| GENDER | 22961.606 | 1 | 22961.606 | 1.108 | .306 | 1.108 | .169 |
| WORKLOAD | 250954.672 | 1 | 250954.672 | 12.115 | .003 | 12.115 | .908 |
| IM | 295108.011 | 2 | 147554.006 | 7.123 | .005 | 14.246 | .883 |
| GENDER * WORKLOAD | 1253.472 | 1 | 1253.472 | .061 | .808 | .061 | .056 |
| GENDER * IM | 2459.878 | 2 | 1229.939 | .059 | .943 | .119 | .058 |
| WORKLOAD * IM | 67680.144 | 2 | 33840.072 | 1.634 | .223 | 3.267 | .299 |
| GENDER * WORKLOAD * IM | 86151.744 | 2 | 43075.872 | 2.079 | .154 | 4.159 | .371 |
| Error | 372864.500 | 18 | 20714.694 | | | | |
| Total | 6242410.000 | 30 | | | | | |
| Corrected Total | 1145144.800 | 29 | | | | | |

a. Computed using alpha = .05

b. R Squared = .674 (Adjusted R Squared = .475)