

センサによって取得したストレスレベルによる 利用者インタラクションの設計・評価

Design and Evaluation of User Interaction using Stress Level Informaion Acquired by Various Sensors

石井 愛弓[†], 飯島 正[‡]
Ayumi ISHII[†], and Tadashi IJIMA[‡]
[†]慶應義塾大学大学院 理工学研究科
[‡]慶應義塾大学 理工学部

[†]Graduate School of Science and Technology, Keio Univ.

[‡]Faculty of Science and Technology, Keio Univ.

要旨

熱画像カメラによる顔温度分布, 心電図による心拍揺らぎ, 近赤外光分光による脳血流, 唾液アミラーゼ活性などの生体情報の計測によってメンタルワークロードを測定することで, 利用者インタラクションの設計に役立てるための, 情報システムのユーザビリティの評価を行う. 従来から主観的評価が優勢であったユーザビリティ評価に対し, 定量的評価を組み合わせる混合アプローチである.

1. はじめに

近年, 情報技術・機器の普及が一層進んでいることに伴い, 多くの業務フローや日常のやりとりが情報システム化されている. しかし高度化したシステムは, 初心者が操作に戸惑うなどの問題が生じることがある. そのため, 誰もが情報システムを適切に使うことができるよう, 利用者インタラクションの一層の向上が求められている.

本研究では, システムの機能だけに目を向けるのではなく, 人とシステムのインタラクションにおけるユーザの生体反応・生体情報に着目しストレスレベルを推定することで, システムの設計・評価に応用する手法について述べる. 従来から主観的評価が優勢であったユーザビリティ評価に対し, 定量的評価を組み合わせる混合アプローチである.

2. 利用者インタラクションの評価手法

利用者インタラクションの向上のための設計手法として人間中心設計が注目されている. その手法では, 利用者状況の理解と明確化, ユーザの要求事項の明確化, デザインによる解決策の作成というプロセスに加えて評価のプロセスがあり, その結果を受けて各プロセスを再度見直す循環型の手法となっている. そのため人間中心設計においては, 評価の適切性が重要となる.

2.1. 既存のユーザビリティ評価手法

表 1: 主な既存のユーザビリティ評価手法と対象ユーザ・利点・欠点

方法	対象	利点	欠点
ヒューリスティック評価法	専門家	個人的なユーザビリティの問題点を見つけられる.	実際のユーザからの発見がない
ユーザビリティ テスト	一般 ユーザ	ユーザの誤解が発見しやすい, 低コスト.	主観的評価になる.
		ユーザの実際の作業状況がわかる.	アポイントメントが難しい.
		ユーザのシステム使用の好みかわかる.	パイロットテストが必要.
		柔軟性があり, 詳しい検証ができる.	時間がかかる. 分析や比較が難しい.
システム使用のログデータ	一般 ユーザ	使用頻度の高い機能がわかる.	分析プログラムには膨大なデータが必要. ユーザのプライバシー侵害

ユーザビリティの評価手法のなかには, 一般ユーザではなくエキスパートを対象

とするものもある [7] が, 本研究におけるユーザビリティ評価の対象は一般ユーザである. 一般ユーザを対象とする既存の評価手法にユーザビリティテストがある (表 1). その手法では, 実際のユーザの作業状況がわかり, システムの問題点の改善につなげることができるが, 一方で, 主観的評価になることや, 分析に時間がかかることが課題となっている.

3. ストレスによるユーザビリティ評価

本研究では, 一般ユーザを対象とした上で, システムを客観的に評価するため, ストレスレベルを利用したユーザビリティ評価の有用性の検討を行う. Nilsen のユーザビリティ概念図 ([9]) では, (1) Learnability(学習しやすさ), (2) Efficiency(効率の良さ), (3) Memorability(記憶しやすさ), (4) Errors(エラー

の起こりにくさ), (5) Satisfaction(満足度) といった尺度が提案されているが, このうち, 近年では, 特に Satisfaction(満足) の評価と重要性が注目されている. Satisfaction は, ユーザがシステムを好きになるか, 楽しく利用できるかといった概念である. いずれの評価尺度にしる, システムを利用する際にユーザがストレスを感じないということにつながっている.

一方, 心理的負担(メンタルワークロード)の主観的評定尺度としてよく知られている SWAT(Subjective Workload Assesment Technique) では, 評価尺度として, (1) 時間切迫性, (2) 精神的努力, (3) 心理的ストレスが挙げられている. また同様に広く知られている NASA-TLX(NASA Task Load Assesment Technique) においては, (1) 知的・知覚的要求, (2) 身体的要求, (3) タイムプレッシャー, (4) 作業成績, (5) 努力, (6) フラストレーションが評価尺度として挙げられている. こうしたストレスを増加させる要素は, 情報システム利用時にも発生しうるものであり, ユーザビリティにおける Satisfaction とストレスとの関連性が示唆される.

本研究では, 従来より主観的評価が優勢であったユーザビリティ評価に定量的評価を組み合わせる混合アプローチの一環として, センサによってユーザの生体情報を取得し, それに基づいたストレスレベルの推定を行うことを試みる.

4. ストレスと生体情報の関連性 (関連研究と予備実験に基づく仮説)

ストレスと生体情報の関連を検討する研究は広く行われている. 例えば, Nis Hjortskov らは, コンピュータ作業中の人物の心電図と血圧を取得し, 心拍変動において低周波成分は変化がなく, 高周波成分(HF)は減少し, また, 血圧は上昇したという結果を報告している [1]. 他にも, 心拍・呼吸・血圧を用いたストレス評価手法や顔面皮膚温とストレスの関連を検討した報告がある [2]-[6]. 本研究では, これらの関連研究と生理学的知識, 予備実験結果をもとに, まず図1の仮説を立て, その仮説の信頼性について検討するため, 必要な生体情報取得システムを構築しながら仮説の検証実験を行った.

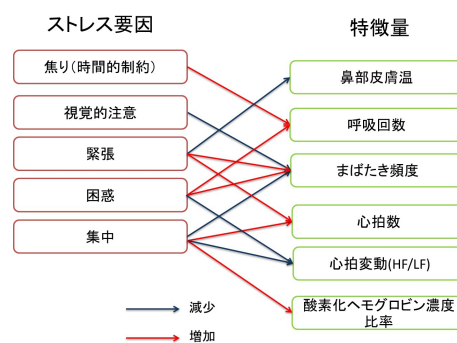


図 1: ストレス要因と特徴量に関する仮説

5. 生体情報取得のためのセンシング技術

ストレスに関わる生体情報として, まばたき(瞬目), 鼻部皮膚温, 呼吸, 心拍(心電図), 脳活動(脳血流), 唾液アミラーゼ活性を取り上げる. ここでは, それら生体情報のセンシングのため, 可視光カメラ, 熱画像カメラ, 加速度センサ, 心電計, NIRS(近赤外光脳血流計), 唾液アミラーゼモニタを使用し, 瞬目判定, 鼻部皮膚温測定, 呼吸検出, 心電データの処理のために各サブシステムを構築した.

5.1. 使用したセンサと生体データ

5.2. 瞬目(まばたき)検出サブシステム

瞬目(まばたき)の検出には, 可視光カメラで被験者の顔を撮影したデータを使用した. 予め眼瞼を閉じた状態の左目のテンプレートを用意しておき, カメラの入力画像とテンプレートマッチングを行う. 探索により閾値を超える類似度が計算された場合に, 瞬目と判断し, 回数をカウントした.



図 2: 瞬目検出

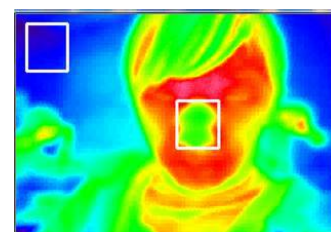


図 3: 鼻部皮膚温測定

5.3. 鼻部皮膚温測定サブシステム

熱画像カメラで被験者の顔を撮影し, 鼻部の平均温度を1秒毎に計算したデータを用いた.

5.4. 呼吸検出サブシステム

息を吸うと鼻孔の温度が減少し, 息を吐くと鼻孔の温度が上昇することから, 鼻孔を熱画像カメラでとらえることで, 呼吸状態を推定した. 時系列で配列に保存された鼻孔温度に対して, 移動平均をとり,

平滑化処理を行った上で、ある一定区間内での局所最大値と局所最小値を抽出した。局所最大値から局所最小値までを吐いている状態、局所最小値から局所最大値までを吸っている状態と判断した。

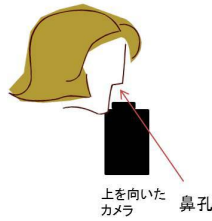


図 4: 鼻孔の撮影の様子

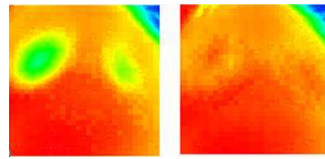


図 5: 鼻孔の画像

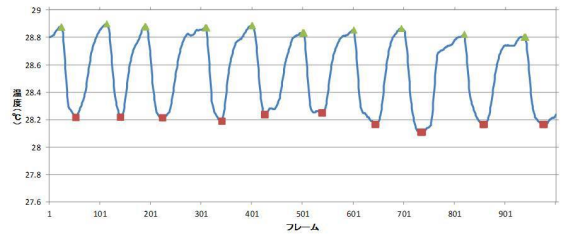


図 6: 鼻孔温度変化

5.5. 心電データ処理サブシステム

心電計で得られた心電図のデータに対し、平滑化処理を行い、ノイズを除去したうえで、一定の閾値を超える範囲での局所最大値を R 波として検出した。この R 波と次の R 波の間隔を RR 間隔とし、心拍数を求めた。また、自己回帰モデルを用いたスペクトル解析を行うことで、低周波成分 (LF : 0.04~0.15) と高周波成分 (HF : 0.15~0.40) に分解し、この LF と HF の比をとることで心拍変動を算出した。

6. ストレス課題と生体情報

ストレス課題と生体情報の関連を検討するために2つの実験を行った (実験 1: 演算スパンテストをストレスラーとする実験と実験 2: フラッシュ暗算をストレスラーとする実験)。いずれも被験者は健康な 20 代男性 10 名である。

6.1. 実験 1: 演算スパンテスト

6.1.1. 実験手順

コンピュータ上で行う演算スパンテストをストレスラーとして、被験者にストレスを負荷し、各生体情報を取得した。演算スパンテストは、ワーキングメモリ容量の個人差を測定する課題であり、演算フェーズとアルファベットの記憶フェーズが交互に表示され、簡単な演算の正誤判断をしながら、計算の合間に提示される単語または数字を記憶する。被験者は、2 分間の待機の後、1 度目は通常の演算スパンテスト、2 回目はラジオを聴きながらの演算スパンテストを続けて行うテスト終了後、再び 2 分間待機し、実験終了となる。

6.1.2. 演算スパンテストによるストレスと生体情報の関連

(a) 鼻部皮膚温

5 名の被験者からデータを取得した結果、以下の傾向がみられた。

- 課題開始前の待機時間に鼻部皮膚温が減少する。
- 課題取組中は温度変化が相対的に少ない。
- 課題終了後から鼻部皮膚温が上昇する。

ある被験者一名の時系列の鼻部皮膚温の変化を図 7 に示す。また、5 名の被験者について、待機中、課題 1 取組中、課題 2 取組中、安静中の前後の温度データの差を図 8 に、各個人のデータを表 9 に示す。

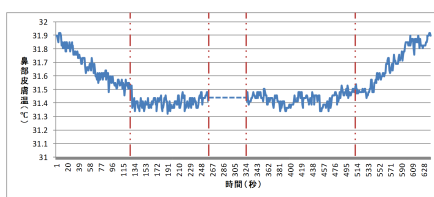


図 7: ある被験者の鼻部皮膚温変化

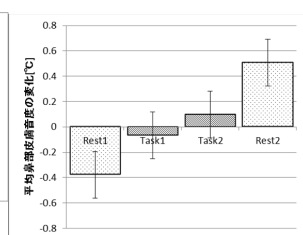


図 8: 鼻部皮膚温度変化

図 9: 鼻部温度変化

	A	B	C	D	E
待機	-0.370	-0.766	-0.259	-0.375	-0.116
課題 1	+0.046	-0.335	-0.023	+0.116	-0.139
課題 2	-0.070	+0.208	-0.277	+0.578	+0.046
安静	+0.416	+1.086	+0.185	+0.647	+0.203

図9より、全ての被験者が、待機中に鼻部温度が低下し、安静中は、上昇していることが確認できる。課題取組中の温度変化は、待機中、安静中と比較して、少ない傾向がみられる。いずれの被験者も待機中に下がった皮膚温を維持しており、温度変化には現れない結果となった。

(b) 瞬目(まばたき)

瞬目検出に成功した7名の被験者のデータのみ使用する。

7名の被験者の瞬目回数を各フェーズ毎にカウントし、単位時間あたりの瞬目回数を時系列順に表したものが図10である。右端は課題終了後の安静時の単位時間あたりの平均回数である。

いずれの被験者も、演算スパンテスト中は、安静時よりも瞬目頻度が減った。さらに、演算スパンテスト中の瞬目についてフェーズ(「暗算と記憶」、「問題の回答」、「小休止」)毎に比較を行った。図10にも示すように、「問題の回答」フェーズでは、他の段階に比べ、まばたき頻度が減る。一方、「小休止」中の瞬目頻度は最も高くなった。

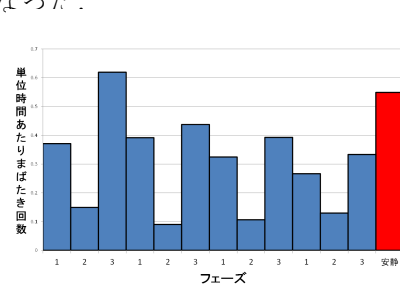


図10: 単位時間あたりの瞬目回数
他の段階に比べ、まばたき頻度が減る。

(c) 呼吸

5名の被験者からデータを取得した。まず安静時と課題取組時の呼吸の長さを比較する。全被験者の安静時と課題取組時の平均呼吸時間を図11にあらわす。図11より、課題取組中は、呼吸が浅く、速くなっていることがうかがえる。

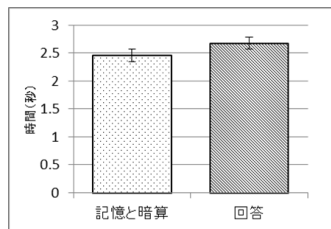


図11: 安静時と課題取組時の平均呼吸時間

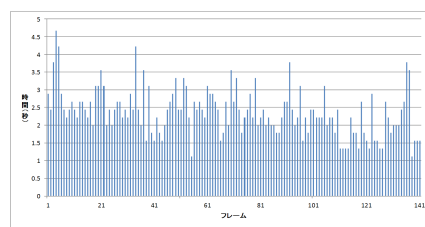


図12: 呼吸の長さ

図12は、ある被験者の課題取組中の呼吸の長さを時系列順にグラフ化したものである。呼吸の長さに注目しても同様に、記憶と暗算中は、呼吸が短く、浅い呼吸になっていることがうかがえる。

6.2. 実験2: フラッシュ暗算

6.2.1. 実験手順

フラッシュ暗算をストレッサーとした。フラッシュ暗算とは、コンピュータ上に連続的に現れる数字を累計し、計算結果を答えるテストである。被験者、4分間の安静の後、まず練習としてフラッシュ暗算を3セットおこなった後、本番としてフラッシュ暗算を10セット連続で行う。終了後、再び4分間安静状態を保ち、実験終了とする。実験1と同様にセンサをセットする。

6.2.2. フラッシュ暗算によるストレスと生体情報の関連

(a) 心拍と心拍変動

心電図は、センサの装着に失敗し正しいデータを得られなかった1名を除く、6名の被験者からデータを取得した。心拍数に関する結果を図13に示す。テストによるストレス負荷時と安静時を比較すると、いずれの被験者もストレス負荷時の方が心拍数が多く、Wilcoxonの符号付順位検定を用いて検定を行った結果、ストレス負荷時の心拍数の有意な上昇が認められた ($p < 0.05$)。

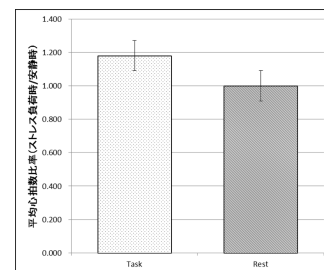


図13: 平均心拍比率変化

次に、心拍数変動について心拍と同様の検定を行った結果、ストレス負荷時のHF、LF、HF-LF、HF/LFのいずれも有意な減少は認められなかった。また、6名の被験者をHF/LFが相対的に高い(0.8以上)安定傾向が比較的低いグループとそれ以外の2群に分け、それぞれの群について、実験中の各フェーズでHF/LFの有意な変化があるかどうか検定を行ったが、有意な差は認められなかった。

(b) 脳活動

図 14 に被験者の一人の NIRS データ例を示す。図 14 では、テスト時にはどちらも酸素化ヘモグロビン濃度 (oxy-Hb) の上昇の傾向があり、脱酸素化ヘモグロビン濃度 (deoxy-Hb) は一定もしくは減少の傾向があることが確認できる。一方、テスト終了後の安静時には、oxy-Hb の減少傾向と deoxy-Hb の上昇傾向が確認できる。

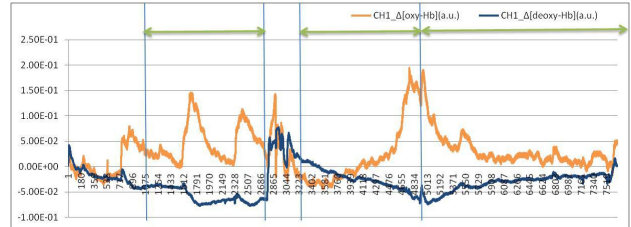


図 14: 被験者の NIRS データの一例

ここで、取得した 5 名の被験者のデータを利用して Wilcoxon の符号付順位検定を用いて検定を行った結果、「安静」と比較した「ストレス負荷」のフェーズでは、deoxy に対して oxy は有意に上昇していると認められた ($p < 0.05$).

(c) 鼻部皮膚温

7 名の被験者からデータを取得した。「安静」と「ストレス負荷」の各フェーズの平均鼻部皮膚温を比較すると、Wilcoxon の符号付順位検定を用いて検定を行った結果、「ストレス負荷」時の鼻部皮膚温の有意な減少が認められた ($p < 0.05$).

(d) 瞬目 (まばたき)

7 名の被験者からデータを取得した。安静時とストレス負荷時の平均瞬目頻度 (1 秒あたりの瞬目回数: 回/秒) を比較すると、6 名の被験者はストレス負荷時の方が瞬目頻度が低く、Wilcoxon の符号付順位検定を用いて検定を行った結果、ストレス負荷時の瞬目頻度の有意な減少が認められた ($p < 0.05$).

6.3. 全体の結果と仮説の比較

実験から得られた結果は以下である。心拍変動に関しては有意な結果は得られなかったが、それ以外の特微量では、仮説を裏付けるものとなった。

- (1) 焦り (時間的制約) による呼吸回数増加
- (2) 視覚的注意によるまばたき頻度減少
- (3) 緊張による鼻部皮膚温減少とまばたき頻度増加
- (4) 困惑による呼吸回数とまばたき頻度増加
- (5) 集中によるまばたき頻度減少と心拍数増加, 酸素化ヘモグロビン濃度比率増加

7. 情報システムのユーザビリティ評価

7.1. 実験 1: 電子メールサービスのユーザビリティ評価

メールサービスに対する複数の操作課題を与え、操作中の様子を観察すると共にセンサでストレスレベルを測定する実験を行った。被験者は 20 代大学生または大学院生 6 名 (コンピュータ一般と他の E メールシステムの利用経験は十分にあるが、当該システムは初めて利用するユーザ) である。

7.1.1. 質問紙による主観的評価

表 2: ユーザビリティ評価のための質問紙法

尺度	内容
CUSI (Kirakowski)	システムに支援されていると感じる程度 (Competence) とシステムを好む程度 (affect)
SUS (Brooke)	頻繁に使用したいか、複雑でないか、簡単に使えるか、使用するのに技術サポートは必要か
SUMI (Kirakowski)	反応は遅いか、ガイドはわかりやすいか、情報は明確か、毎日使いたいか
QUIS (Shneiderman)	システム全体の使用感、画面、用語とシステム情報、学習、システムの能力
WAMMI (Kirakowski,Claridge,Whitehand)	Attractiveness,Control,Efficiency, Helpfulness,Learnability
Web Usability Scale (イード・富士通)	好感度、役立ち感、信頼性、操作のわかりやすさ、構成のわかりやすさ、見やすさ、反応の良さ

センサによる生体情報の取得に加え、質問紙によって被験者自身の主観的評価を得た。表のような質問紙法が知られている [7] が、今回は Shneiderman の QUIS (Questionnaire for User Interadction Satisfaction) を参考に、今回の実験用に一部改変したものを使用した。QUIS の評価シートには、評価対象のシステムやコンピュータ経験の利用経験に関する質問の後に、「システム全体の使用感」、「画面」、「用語とシステム情報」、「学習」、「システムの能力」といった各尺度について、それぞれに 8 段階+1 (未適用の場合) で回答する。QUIS は Chin et al. に 150 人を対象にした結果が掲載されており、心理尺度として適切と言えることが報告されている。

7.1.2. 実験に用いた情報システムにおける既知の問題点

実験に用いた電子メールサービスでは、事前調査とヒアリングによって実験前から幾つかの問題点が指摘されている。その問題点とそれによって影響をうけると考えられる QUIS の尺度を表 3 に示す。

本研究では、これらの既知の問題点を含むページの

利用をイベントとしてとらえる。各イベント期間中と安静時と比較し、各生体情報の変化からストレスレベル変化の推定を行う。それに基づいて表 4 に挙げた既知の問題点を検出できるかどうかを検討する。

7.1.3. 実験結果

6名の被験者を対象とした。表 4 では、イベント時と安静時の生体情報を比較し、Wilcoxon の符号付順位検定を行った結果、 $p < 0.05$ で有意な変化があるといえるものに、◎、 $p < 0.10$ で有意な変化があるといえるものに、○を与えている。

なおイベント時に見られた変化の傾向は、瞬目の減少・NIRS(酸素化ヘモグロビンと脱酸素化ヘモグロビンの差)の上昇・心拍数の上昇・心拍変動(LF/HF)の上昇であった。

また、図 18 に表わすように、情報システムの負荷がどのようにパラメータに影響与えるのか、因子分析を行った。これをもとに、各パラメータに重み付けをして情報システムの負担の大きさを表現すると次式となる：
 情報システムによる負荷のレベル = $0.214 * \text{心拍数増加量 (回/分)} + 0.556 * \text{心拍変動増加量 (LF/HF)} + \text{脳血流ヘモグロビン変化量 (oxyHb-deoxyHb)} + \text{まばたき減少量 (回/秒)}$ 。

7.2. 実験 2(今後の予定): Web 履修登録システムの改善に向けて

大学等で使われている Web 履修登録システムの簡略版を評価用に作成した。本手法によるユーザビリティ評価によって改善すべき点を発見し、よりユーザビリティの高いインタラクションを選択することで、全体として満足度(Satisfaction)の高いインタラクション設計を目指す。

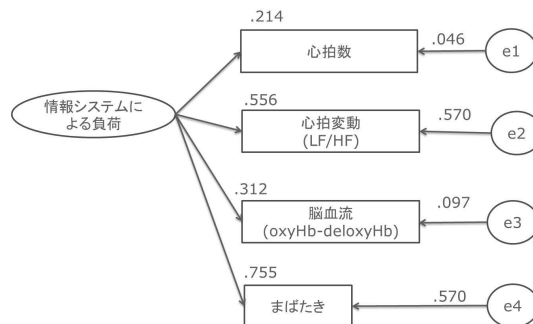


図 15: 情報システムと生体情報



図 16: 構築した履修申告システム

表 3: 実験用メールサービスの既知の問題点

問題点	影響を及ぼす要素
10 文字以上の入力	画面、学習
フォームの大きさが小さすぎる	画面
入力フォーマットが不明確	用語とシステム情報
使いたいページの場所がわからない	用語とシステム情報
表示順序が不明確	学習、用語とシステム情報
システムの速度が遅い	システムの能力
間違えると途中から修正できない	システムの能力
予想外の画面へ遷移	システムの能力

表 4: 被験者のイベントに対する各生体情報の変化の有無

イベント ID	問題点	瞬目	NIRS	心拍	心拍変動
1	10 文字以上の入力 フォームの大きさが小さすぎる 入力フォーマットが不明確	◎	○		
2	使いたいページの場所がわからない システムの速度が遅い	○	◎		◎
3	(待ち時間の発生)				
4	間違えると途中から修正できない (手戻りの発生)	◎	◎	◎	
5	予想外の画面へ遷移	◎	◎	◎	

参考文献

- [1] Nis Hjortskov, Dag Rissen, Anne Katrine Blangsted, Nils Fallentin, Ulf Lundberg, and Karen Sogaard: "The Effect of Mental Stress on Heart Rate Variability and Blood Pressure during Computer Work," pp.84-89, 2004.
- [2] Kirschbaum C, Pirke K, -M, Hellhammer D, H: "The 'Trier Social Stress Test' ? A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting," Neuropsychobiology, No.28, pp.76-81, 1993.
- [3] Ushiyama, K., Ogawa, T., Ishii, M., Ajisaka, R., Sugishita, Y., Ito, I.: "Physiologic neuroendocrine arousal by mental arithmetic stress test in healthy subjects," The American Journal of Cardiology, No.67, pp.101-103, 1991.
- [4] Seraganian, P., Szabo, A., and Brown, T. G.: "The Effect of Vocalization on the Heart Rate Response to Mental Arithmetic," Physiology & Behavior, No.62, pp.221-224, 1997.
- [5] 下野太海, 大須賀美恵子, 寺下裕美: "心拍・呼吸・血圧を用いた緊張・単調作業ストレスの評価手法の検討," 人間工学, Vol.34 No.3, 1998.
- [6] 岩田洋夫: "サーモグラフィを利用した知的作業の定量評価," 計測自動制御学会論文集, No.24, pp.107-111, 1988.
- [7] 黒須正明, 他: "人間中心設計の基礎," 近代科学社, 2013.
- [8] 原田悦子, 篠原一光: "注意と安全," 北大路書房, 2011.
- [9] ヤコブニールセン, 三好かおる(訳): "ユーザビリティエンジニアリング原論," 東京電機大学出版局, 2002.