

## ユーザビリティ評価における集中と作業成績の関連性の調査 An Investigation of a Relationship between Concentration and Task Performance of a Usability Evaluation

宮崎雄宇<sup>†</sup> 小俣昌樹<sup>‡</sup> 茅暁陽<sup>‡</sup>  
Yu Miyazaki<sup>†</sup> Masaki Omata<sup>‡</sup> Xiaoyang Mao<sup>‡</sup>

### 1. はじめに

従来行われているユーザビリティ評価では、作業成績（タスクを終えるまでの時間やどのくらい達成できたか）を測定し、ユーザの状態（製品に対する満足度、気分、感情など）を調査する。つまり、ユーザの状態と作業成績は独立したものと考えられている。

一方でユーザの状態が作業成績に影響するという報告もある。阿部ら[1]は、計算問題の副次タスクを課すことで意図的に漫然状態を作りだし、ドライビングシミュレータにおける信号に対する反応時間に変化があるのかを調査した。実験結果から、ブレーキ時の反応時間に変化が出ることを報告した。また門間ら[2]は、歌詞を含む音楽を聞く場合と聞かない場合で文章課題を行い、作業数や誤答率に変化があるのかを調査した。その結果、歌詞を含む音楽を聞く方場合は、聞かない場合に比べて誤答率が増加していることがわかった。このように、もしユーザのそのときの状態によって、作業成績が変わってしまうのなら正確なユーザビリティ評価は難しい。

そこで、本研究はユーザの状態の中でも集中に注目し、1つのシステムを提案する。そのシステムの考えを図1に示す。とある作業を完了させるには通常 10 分かかるとする。ユーザが集中を保っているのなら同じように 10 分で終わられるが、集中が保てていないのなら 10 分より多くの時間がかかってしまう。このとき、集中が保てていないその分、時間を修正するようなシステムである。

ユーザが集中しているか否かは、生体信号によって推定する。生体信号を使用する利点として「質問紙調査などの主観評価と違い客観性を持っている」「自動的に取得できる」がある。そのため、生体信号からユーザの状態を推定する研究[3][4]、推定したユーザの状態をユーザビリティ評価に利用する研究[5]が行われている。

本研究ではタスクによって集中を変化させたときの、作業成績と生体信号の関連性を調査していく。しかし、集中と作業成績の関連性を調べた研究は少なく、その関連性は明らかにされていない。本研究で提案する図1のようなシステムを開発するには、まずこの関連性をより明らかにする必要がある。

### 2. 集中と作業成績の関連性

本研究で扱う「集中」とは、精神の集中であり、作業

<sup>†</sup> 山梨大学大学院医学工学総合教育部修士課程  
コンピュータ・メディア工学専攻, Department of  
Education Interdisciplinary Graduate School of Medicine and  
Engineering, University of Yamanashi

<sup>‡</sup> 山梨大学大学院医学工学総合研究部, Interdisciplinary  
Graduate School of Medicine and Engineering, University of  
Yamanashi

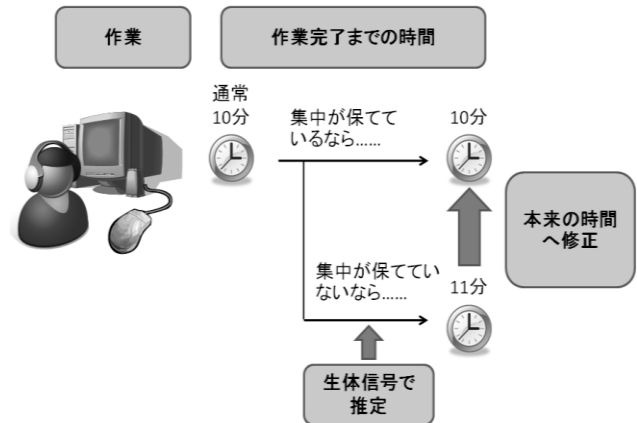


図1 集中により変化するタスク完遂の時間

に対してのものである。本研究ではタスク中に聴覚刺激を与えることで集中を乱れさせ、作業成績を低下させる。使用する聴覚刺激は TBS radio podcasting の「柳瀬博一・Terminal」というラジオである。内容としては、「自動車産業」「仕事」「フェイスブック」などのテーマに沿って二人以上の人間が会話するというものである。音は耳から自動的に入ってくるものであり、実験タスク自体を邪魔することは無い。また、音楽ではなくラジオを使用している理由として、ラジオは話が展開していくことで被験者がつい話の内容に気を取られてしまい、より集中を低下させると考えたからである。

実験を2つ行う。ひとつ目は、細かな単純作業である、ディスプレイ上のエリアを素早くクリックする実験である。2つ目は精神作業である計算課題を行う実験である。脳が行う心の活動（精神活動）をともなった作業が精神作業であり、言語を理解したり、計算を行ったりすることがこれに当たる[2]。志水ら[6]は、BGM が計算課題の正答数に影響しないと報告しているが、本研究では音楽ではなくラジオであり、計算手法も大きく異なる。志水ら[6]の計算問題は複雑な問題ではあるが筆算を使わせており、あまり脳に負担をかけず機械的に解けてしまう。そこで本研究では、計算を解くことに集中が必要となるように、難易度の高い問題を暗算で解く実験を行う。

### 3. ターゲット選択における作業成績の変化の検証

本実験の目的は、ラジオによって集中を変化させたときの作業成績と生体信号の関連性を調査することである。本実験ではスピードタスクとリズムタスクの2種類を行い、スピードタスクは「細かな作業を素早く正確に行う」、リズムタスクは「細かな作業をリズムよく正確に行う」ことが求められる。また2つのタスクに関して、十分な訓練を行った上でタスクを行う。

### 3.1 実験環境と測定する生体信号

図 2 に実験の様子を示す。被験者の正面に 17 インチディスプレイを設置し、ラジオを聞くためにヘッドホンを使用する。

測定する生体信号は、皮膚伝導率、脈波、呼吸である。また、脈波からは脈拍数、HF(副交感神経指標)、LF/HF(交感神経指標)、呼吸からは呼吸数を算出する。これらの生体信号は緊張や集中などのストレス指標として使用されている。生体信号には個人差があるため、絶対値による比較は有用ではない。そこで、タスクを行ったときのデータから個人ごとに平均値を算出し、標準化を行う。測定には、Thought Technology 社の ProComp Infiniti™を使用する。本実験では、左手人差し指と薬指に皮膚伝導センサ、左手親指に脈波センサを、腹部に呼吸センサを装着する。

### 3.2 実験手順：スピードタスク

このタスクは、ディスプレイに表示された 2 つのエリアを交互にクリックしていくものである。図 3 のような画面がディスプレイ上に表示される。左右両端には、大きな四角と小さな四角が描かれており、小さな四角の部分を Score エリア、その外側の部分を Error エリアとする。中央にはタスクを開始する Start ボタン、残り時間を表示する〈Time〉、得点を表示する〈スコア〉、〈エラー〉がある。

Start ボタンを押すことでタスクが開始される。まずは左側から始まり、左右の Score エリアを交互にクリックしていく。このとき被験者には出来るだけ素早く、正確にクリックすることを心がけてもらう。Score エリアをクリックするたび、〈スコア〉に 1 点加算される。ただし、Score エリアを外れて Error エリアをクリックしてしまうと、〈エラー〉に 1 点加算される。〈エラー〉は成績としてマイナスであるため、Score エリアを正確にクリックすることが求められる。この左右交互に Score エリアをクリックしていくタスクを、制限時間 3 分で行ってもらう。

Score エリアの大きさは 3 サイズ (大: 60×60, 中: 30×30, 小: 10×10) 用意し、ラジオを聞く場合と聞かない場合の 2 種類のタスクがある。すなわち、3 サイズ×2 種類=6 回のタスクとなり、タスクの合計時間は 18 分である。

### 3.3 実験手順：リズムタスク

スピードタスクと同じで、左右の Score エリアをクリックしていくタスクである。違う点は、1 秒ごとに左右のエリアの表示非表示が切り替わる。左のエリアが現れているときは右のエリアは非表示となり、右のエリアが現れているときには左のエリアは非表示となる。表示されているときだけ、エリアをクリックできるようになっている。すなわち、1 秒ごとに切り替わる左右のボタンをリズムよく正確に押していくことが求められる。

ただし、リズムタスクでは必ずしも左右交互に押す必要は無い。リズムタスクは難易度が高く、1 秒の間にエリアをクリックできないことが起こりえるため、片方のエリアを連続して押しても〈スコア〉もしくは〈エラー〉に加算されるようになっている。ただし、エリアを押さ



図 2 実験の様子

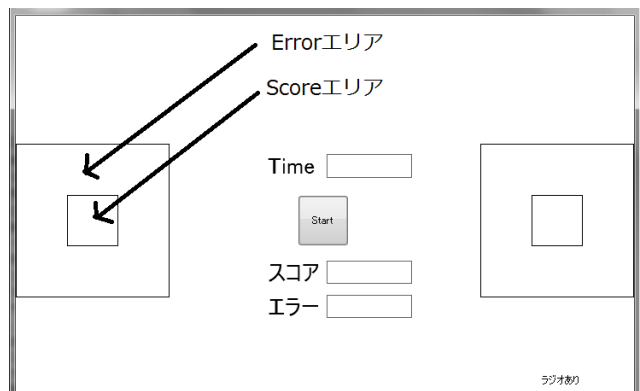


図 3 実験画面

ずに見逃してしまった場合は 1 回見逃すごとに、〈ミス〉に 1 点加算される。これはエラーと同様に成績としてマイナスであるため、出来るだけ見逃さないことが求められる。Score エリアのサイズや制限時間はスピードタスクと同じである。

### 3.4 実験結果

#### 3.4.1 作業成績への影響

被験者は 22 歳から 26 歳の 8 人 (男性 7 人, 女性 1 人) である。スピードタスクの作業成績は-(エラー)得点、リズムタスクの作業成績は-(エラー+ミス)得点とする。本来マイナスの成績であるエラーやミスを作業成績としているのは理由がある。2 つのタスクにおいてクリックで得点を得られる範囲に大きな差があり、スピードタスクでは被験者次第でいくらかでも点数があがるのに対し、リズムタスクは上げられる点数が決まっている。一方、マイナスの成績は、自分がエリアをクリックしようとして失敗した数であり、2 つのタスクで意味合いが同じとなる。そのため、あえてマイナスの成績を作業成績としている。被験者全員の作業成績の合計を図 4、図 5 に示す。それぞれ難易度、ラジオを要因とした分散分析を行ったところ、スピードタスクにおいてサイズの大きさで有意差( $p < 0.01$ )があり、多重比較ではサイズ小とサイズ中、サイズ小とサイズ大に有意差があった。ラジオでは有意差は見られなかった。また、リズムタスクも同様であり、サイズの大きさで有意差( $p < 0.01$ )があり、多重比較ではサイズ小とサイズ中、サイズ小とサイズ大に有意差があった。ラジ

オでは有意差は見られなかった。どちらも、サイズが小さくなるとマイナスの得点が増え、作業成績が下がるという結果となった。

### 3.4.2 生体信号への影響

測定した生体信号から、皮膚伝導(平均)、皮膚伝導(標準偏差)、脈拍数、HF、LF/HF、呼吸数を計算し、難易度、ラジオを要因とした分散分析を行った。その結果、リズムタスクにおける皮膚伝導平均がラジオの有無で有意差が見られ、ラジオを聞くことによって上昇した( $p<0.05$ )。他の項目に有意差は見られなかった。

### 3.5 考察

実験の結果、ラジオを聞いた場合に作業成績が良くなる被験者と悪くなる被験者の両方がおり、本実験では「ラジオによって集中を乱す」ことは出来なかった。質問紙調査では「ラジオによって集中が乱された」という意見がある一方で「単純作業のみで眠くなってしまい、ラジオがある方が集中できた」という意見もあった。実験結果から、ラジオを聞くことで皮膚伝導平均、つまり覚せい度が上昇していることがわかる( $p<0.05$ )。元々眠気がある人がラジオを聞くことで覚せいし、タスクに集中できるようになったとも考えられる。このように、作業成績は人の眠気にも左右される可能性があるため、本実験で行ったような単純作業は、作業成績を測るためには適切ではなかったと言える。もしくは、単純作業でもタスク時間を短くするなどして、眠気が起こらないような工夫が必要である。

## 4. 計算問題における作業成績の変化の検証

本実験の目的は、精神作業である計算問題をタスクとして、前述の実験と同様による作業成績と生体信号の関連性を調査することである。精神作業とは脳が行う心の活動(精神活動)をともなった作業であり、前述の実験であったエリアをクリックしていくタスクとは明確な違いがある。

### 4.1 実験手順

図 6 のように、3つの数字を加算する計算問題を行った。計算はすべて暗算で行い、答えの数値のみを紙に記入する。制限時間5分の中で、素早く正確に解くことが目標となる。難易度は2種類用意し、加算する数字が2桁と3桁の場合である。また、計算を行いながらラジオを聞く場合と聞かない場合の2種類のタスクを行った。タスク回数は4回でのタスク合計時間は20分となる。また、タスク間で1分間の休憩を挟んだ。

### 4.2 実験結果

#### 4.2.1 作業成績への影響

被験者は21歳から23歳の7人(男性6人、女性1人)である。計算問題を行った結果から、被験者それぞれの作業成績(作業数、正答数、正答率)を調べた。この作業成績を、難易度とラジオを要因とした分散分析を行った結果、作業数では難易度で有意差( $p<0.01$ )が確認され、難易度の低い場合に作業数が多くなった。また、ラジオの要因では有意傾向( $p<0.1$ )が見られ、ラジオがない場合に

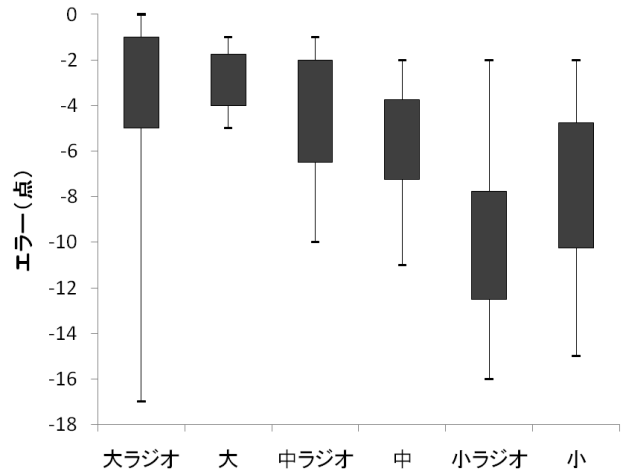


図 4 スピードタスクにおける作業成績

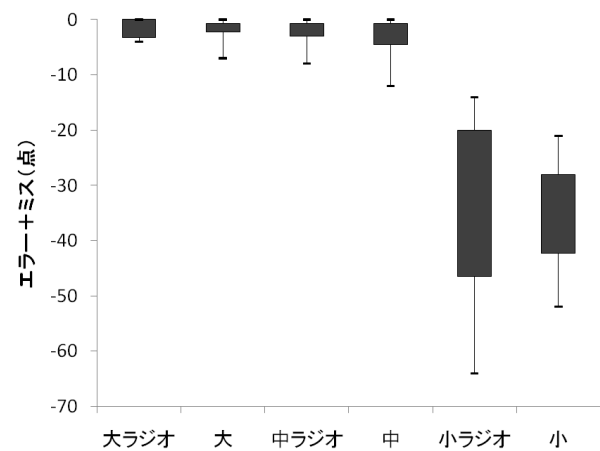


図 5 リズムタスクにおける作業成績

22 + 37 + 81 =	376 + 474 + 372 =
41 + 36 + 65 =	328 + 656 + 942 =
29 + 39 + 98 =	696 + 968 + 391 =
.	.
.	.

図 6 計算問題例

作業数は多くなる傾向が見られた。正答数を図 7 に示す。正答数では難易度で有意差( $p<0.01$ )が確認され、難易度が低い場合に正答数が多くなった。また、ラジオの要因で有意差( $p<0.05$ )が確認され、ラジオを聞かない場合にも正答数が多くなった。また、交互作用における単純主効果が確認され( $p<0.05$ )、ラジオ有りでの難易度の違い、ラジオ無しでの難易度の違い、2桁の計算問題におけるラジオの有無でそれぞれ有意差( $p<0.01$ )が確認された。正答率では難易度で有意差( $p<0.05$ )が見られ、難易度が低い場合に正答率が高かった。ラジオの要因での有意差は見られなかった。以上の結果から、2桁の計算問題と3桁の計算問題には難易度の違いがあり、ラジオによって集中が妨げられていることが確認できた。

#### 4.2.2 生体信号への影響

それぞれの生体信号で、難易度とラジオを要因とした分散分析を行った結果、皮膚伝導の平均値では難易度で有意差( $p<0.01$ )が見られ、難易度の高い方が、値が高くなった。一方でラジオでの有意差は見られなかった。皮膚伝導の標準偏差(タスク内でのばらつき)では難易度で有意差は見られなかった。ラジオでは有意差( $p<0.01$ )が見られ、ラジオがない場合に値が高くなった。HF では難易度、ラジオともに有意差は見られなかったが、交互作用( $p<0.05$ )が見られた。下位検定を行うと、2桁の計算におけるラジオの有無で単純主効果( $p<0.1$ )が見られ、ラジオ有りの場合に HF の値が高くなった。脈拍数、LF/HF、呼吸数では難易度、ラジオともに有意差は見られなかった。

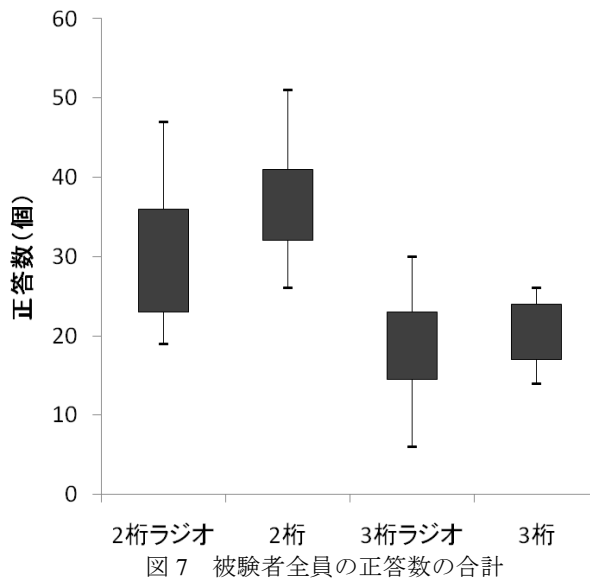
#### 4.3 考察

難易度を要因とした場合、作業数、正答数、正答率ともに有意差が見られ、2桁の計算問題と3桁の計算問題では明確に難易度の差があることが判明した。またラジオを要因とした場合、作業数で有意傾向( $p<0.1$ )、正答数で有意差( $p<0.01$ )が認められた。また質問紙調査でも多くの被験者が「ラジオによって集中が乱された」と回答していた。このことから、ラジオによって集中を妨げられ、作業スピードの低下や誤答をまねいたと考えられる。また、正答数では2桁の計算問題におけるラジオの有無で単純主効果が確認され、3桁の計算問題では確認されなかったことから、難易度の間でも差が出た。志水ら[6]の実験でもBGMを聞かせても計算問題の正答数に影響は見られなかった。このことから計算作業(精神作業)だから影響が出る、出ないという話ではないことがわかる。今後様々な計算問題で実験を行い、詳しく調べていく必要がある。

生体信号に関して、皮膚伝導の平均値は、難易度が高い場合に高くなった。これは、計算が難しくなることで緊張が増したと考えられる。また、皮膚伝導の標準偏差(タスク内でのばらつき)はラジオがない場合に高くなった。皮膚伝導の標準偏差(タスク内でのばらつき)が高い、つまり覚せい度のばらつきが多いのは、計算問題1問を解く過程で集中に大きな差が出ている可能性があるが、はっきりとはしていない。今後質問紙調査でユーザの状態を詳しく調査し、生体信号と対応づけていく。また、HFでは交互作用が見られ( $p<0.05$ )、2桁の計算におけるラジオの有る場合はない場合に比べてHFが高くなった。これは被験者がラジオに気を取られ、計算問題への集中が妨げられたことがHFに影響を与えたと考えられる。

#### 5. おわりに

本研究では、集中を変化させたときの作業成績と生体信号の関連性を調べるために実験を行った。ひとつめの実験は単純作業のターゲット選択タスク、2つめは精神作業である計算問題をタスクとし、ラジオで集中を変化させることで作業成績への影響を調べた。この結果、ひとつめの実験で行った単純作業には影響が無かったが、2つめの実験で行った精神作業では影響があり、ラジオによって集中が乱されていることが確認された。質問紙調査から、エリアをクリックするだけのような単純作業の場合は眠気を誘発してしまうことがわかり、「ラジオによって集中を乱す」ことが出来なかった。単純作業で作業



成績を測定するときは、タスク時間を短くするなどの工夫が必要である。

生体信号は、難易度が高くなると皮膚伝導の平均値が上昇し、ラジオを聞くことでHFが上昇した。これは、難易度によって緊張し、ラジオを聞くことで集中が妨げられた状態である。以上の結果から、本研究で扱った「集中」を推定するには、皮膚伝導や脈波が有効であると言える。

本実験で集中と作業成績の関連性はより明確になったが、作業成績が変化するときの生体信号の詳細はまだ明らかになっていない。今後質問紙調査でユーザの状態を詳しく調査し、生体信号と対応づけていく必要がある。

#### 参考文献

- [1] 阿部 喜, 宮武 秀樹, 小栗 宏次: 暗算タスクによる漫然運転の誘発と生体信号を用いた評価の試み; 電子情報通信学会論文誌, 基礎・境界 J91-A(1), pp.87-94, (2008).
- [2] 門間 政亮, 本多 薫: 音楽に含まれる言語情報が文章課題に与える影響に関する検討; 人間工学 Vol. 45, No. 3 pp.170-172, (2009)
- [3] Mina, M., Khaled, El., Rana, E. K., James, C., John, J.B. A.: Emotion Detection using Noisy EEG Data, AH'10, No.7, (2010)
- [4] Richard, L. H.: Measuring Emotional Valence during Interactive Experiences: Boys at Video Game Play, CHI'06, pp.1023-1026, (2006)
- [5] 上野秀剛, 石田響子, 松田侑子, 福嶋祥太, 中道上, 大平雅雄, 松本健一, 岡田保紀: 脳波を利用したソフトウェアユーザビリティの評価-異なるバージョン間における周波数成分の比較; ヒューマンインタフェース学会論文誌 10(2), pp.233-242, (2008)
- [6] 志水佳和, 菅千索: 計算課題の遂行に及ぼすBGMの影響について(2); 和歌山大学教育学部教育実践総合センター紀要, 14, pp.103-112, (2004)