

図1. 自由度量度の低さと相関した 前帯状回腹側部の活性

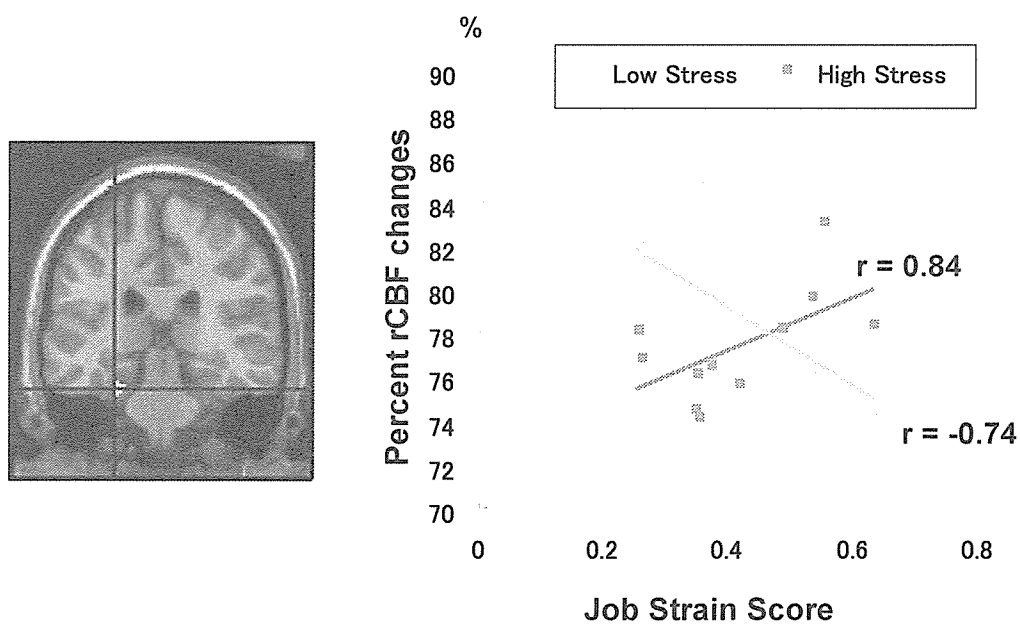


図2. 自覚ストレスと職場ストレインとの交互作用を示した左海馬傍回活性

平成 18 年度厚生労働科学研究労働安全衛生総合研究事業
「過重労働等による労働者のストレス負荷の評価に関する研究」
分担研究報告書

ストレス課題を用いたストレス負荷評価システムの開発：
PET・脳波・心臓血管系活動の同時計測による脳-身体の機能的関連の検討
—確率学習課題による検討—

分担研究者 大平英樹 名古屋大学大学院環境学研究科・助教授

本研究は、慢性的なストレス負荷が心身に及ぼす影響を客観的に評価する方法を確立することを目的としている。そのために妥当な課題を開発し、それに関する基礎的知見を得ることが必要である。本年度は、急性ストレス課題として確率学習課題を用いた検討を行った。健常者に、報酬・罰の随伴性を70%と50%に操作したオペラント学習型の確率学習を遂行させ、¹⁵O 水を用いた陽電子放射断層撮影法(PET)により脳の局所的血流量を測定し、同時に、脳波(事象関連電位)、心臓血管系活動(心拍、血圧)を計測し、各指標間の関連を検討した。その結果、以下の知見を得た。①確率学習課題により、心拍、血圧の上昇を特徴とする典型的なストレス反応を惹起させることができ、この課題が急性ストレス課題として妥当であることが確認された。②課題遂行中に、前頭眼窩野(BA11)、内側前頭前野(BA32、24)、淡蒼球の賦活を認めたが、それは学習が困難な条件でより顕著であった。③学習成績は、事象関連電位のP3成分と心拍出量と関連していた。同時に、P3成分や心臓活動は、課題中に賦活がみられた上記の脳部位の活動と相関していた。これらの結果は、前頭前野領域と基底核の神経ネットワークによりストレス刺激の再評価や対処可能性が表象され、それが行動的対処を規定していく様相を反映しているものと推測される。

A. はじめに

過重労働などの職業性ストレス要因により、心身の疾患を生じる労働者が増加していると言われている。しかしながら、長時間労働などの過重労働が、どのように各種の疾患を生じさせるのかについての生物学的メカニズムについては、いまだ十分に理解されているとは言いがたい。過重労働の健康・疾患への影響を理解し評価するには、慢性ストレス負荷に伴い、脳と身体の機能的関連がどのように変動するかを客観的に評価できる方法が必要とされる。しかし現在までのところ、そうした研究はほとんど進んでいない。そこで本研究は、急性ストレス課題を負荷し、その時点における脳活動と身体生理反応を測定することによって、その反応性から慢性ストレスを推定する方法を開発することを最終的な目的としている。

本年度は、3年研究の2年目として、上記の急性ストレス課題に関する基礎的検討を行う。実験的に急性ストレスを負荷する課題としては、暗算、ストループ・テストなどの認知的課題、人前でのスピーチなどの社会的課題、痛みや冷却などの身体的課題、など、さまざまなものが用いられてきた。本研究では、①陽電子放射断層検査法(positron emission tomography: PET)や機能的磁気共鳴画像法(functional magnetic resonance imaging: fMRI)などの神経イメージングでの脳活動観測に使用可能なこと、②標準化が可能で一定の急性ストレスを持続的に負荷可能で、負荷強度を操作可能なこと、③年齢、性別などの個人差要因に過度に影響されな

いこと、④倫理性に配慮して侵襲性が低いこと、などの要件を考慮して、認知的課題を用いる。

こうした課題として、昨年度の暗算課題(Isowa et al, 2006)についての検討に引き続き、本年度は確率学習課題(stochastic learning task)について検討する。この課題は一種のギャンブル課題であり、被験者は提示される2つの刺激のうち、一方を選択することを求められる。それぞれの刺激がもたらす金銭的報酬と金銭的罰は確率的に定められており、それに従って報酬や罰の提示がなされる。被験者は当初はこのルールを知らされておらず、試行錯誤的な学習により、適切な反応選択を獲得することになる。こうした学習は、心理学ではオペラント条件づけ(operant conditioning)と呼ばれ、長い研究の歴史を有する。急性ストレス課題としての利点は次のとおりである。①時間圧を負荷することにより、暗算課題やストループ課題と同程度のストレス反応を惹起できる(Kimura et al, in press.)。②報酬・罰の確率操作が容易で、ストレスへの対処可能性などの心理的変数を操作しやすい。③行動の選択、意思決定の過程を含むので、学習成績を数量的に評価することができ、ストレス負荷により行動がどのように変容するかの検討が可能である。④近年の認知神経科学的研究により、この課題の遂行時には前頭前野領域、辺縁系、線条体が活動することが明らかになりつつあり(O'Doherty, Dolan, 2006)、神経基盤の理解が容易であると期待される。

本研究では、基礎的検討として報酬・罰の随伴性

を70%と50%に操作した確率学習課題を健常大学生被験者に遂行させ、その際の、脳活動をPETにより測定し、同時に脳波と心臓血管系活動の指標である心拍、血圧をより心臓血管系活動を、我々が既に確立した技法により(Ohira et al, 2006)、それぞれ同時計測する。これによって、報酬・罰による随伴性の評価、それへの対処を担う神経基盤を明らかにすると共に、そうした神経活動が急性ストレス負荷時に末梢身体や意思決定の過程に影響する様相を検討する。

この研究は多施設共同研究として行われ、そのメンバーは、松永昌宏(名古屋大学・愛知医科大学)、磯和勅子(名古屋大学・三重県立看護大学)、野村理朗(東海女子大学)、市川奈穂(名古屋大学)、木村健太(名古屋大学)、村上裕樹(名古屋大学)、金山範明(名古屋大学)、福山誠介(木沢記念病院・中部療護センター)、篠田淳(木沢記念病院・中部療護センター)、山田實宏(木沢記念病院・中部療護センター)、であった。

B. 研究の方法

1. 対象

健常な男子大学生16名(20-24歳)。いずれも、右効きで、正常な視力あるいは矯正視力を有していた。精神疾患、神経疾患の罹患歴のある者は除外した。いずれの被験者も署名をもって実験参加に同意し。本研究は、木沢記念病院・中部療護センターの倫理委員会によって承認された。

2. 方法

1) 課題

2つの、無意味な線画図形がコンピュータ・ディスプレイに呈示され、被験者はそのひとつを700ms以内にキー押しによって選択することを求められた。その3500ms後、100円の金銭的報酬か、-100円の金銭的罰がフィードバックとして与えられた(図1)。これを1試行とし、4分間のブロックで40試行を行った。実験全体では、合計8ブロックが設定された。被験者には、実験参加への謝礼として15000円が支払われたが、確率学習課題の成績に従い、謝金額を増減すると教示した。

2) 随伴性の操作: 実験条件

報酬・罰の随伴性を操作するために、次の3つの実験条件が被験者内(within-subject)要因として設定され操作された。

①70%強化条件

2つの刺激のうち、一方は70%の確率で報酬を、30%の確率で罰をもたらす。他方は逆に、30%の確率で報酬を、70%の確率で罰をもたらす。どちらの図形が高い報酬確率をもたらすかは、被験者間でカウンタ・バランスされた。この条件は、一連の課題として3ブロック、合計120試行、遂行された。

②50%強化条件

いずれの刺激も同様に、50%の確率で報酬、50%の確率で罰が与えられた。この条件は、一連の課題として3ブロック、合計120試行、遂行された。

③統制条件

上記2つの条件と同様な課題を遂行するが、各試行でどちらの刺激を選択するかは、コンピュータにより指示された。つまり、被験者は70%強化条件、50%強化条件と同様な刺激を見て、同様なキー押し反応を行う。統制条件は2ブロック、合計80試行、遂行された。そのうちひとつのブロックでは、報酬・罰の随伴性は70%強化条件と同様とされた。他方のブロックでは、随伴性は50%強化条件と同様とされた。それゆえ被験者は、反応の選択は全く受動的ながら、刺激と報酬・罰間の関連性の学習は可能である。つまり統制条件は、上記2つの条件と比較した場合、意思決定の過程だけを欠いていることになる。統制条件は、PET画像の減算解析のために設定された。

3) 手続き

8ブロックのうち、統制条件は第1ブロックと、第5ブロックに置かれた。第2-4ブロックと、第6-8ブロックにおいて、70%強化条件と50%強化条件が、それぞれ連続的な課題として遂行された。70%強化条件と50%強化条件の順序は、被験者間でカウンタ・バランスされた。統制条件における報酬・罰の随伴性は、それに続く3ブロックと常に同様とした(図2)。これは随伴性の学習を助け、被験者の学習成績の個人差を極端に大きくしないためにも寄与する。

被験者は、この課題がギャンブルであるとの教示だけを受け、随伴性などについては一切の教示を受けなかった。練習の後、1ブロック4分の暗算課題を、11分の間隔をおいて、8ブロック連続的に遂行した。各ブロックにおいて、課題開始1分の時点からPETによる局所脳血流(regional cerebral blood flow: rCBF)の測定が行われた。実験期間を通じて、脳波と、心臓血管系指標として心拍と血圧を連続的に測定した。

4) PETによる神経イメージング

各ブロックの開始と共に、被験者の肘静脈から370MBqの¹⁵O水を30秒注入し、続く1分間、3DモードによるPETスキャンが行われた。スライス厚4.5mm、2mm x 2mmの空間解像度で35スライスの画像を得た。

5) 生理的指標の測定

①脳波

Fz, Cz, Pzの頭皮上3部位に銀・塩化銀電極を装着し、鼻尖を基準として脳波を導出し計測した。生体アースは前額部に装着された。各電極の抵抗は5 μ Ω 以下に保たれた。

②心臓血管系指標

四肢誘導法による心電図(ECG)をMP100システム(BIOPAC社製)により観測し、R波を検出して心拍率(HR)を求めた。非侵襲的指尖血圧計(ポータプレス

2: TNO バイオメディカル社製)により測定し、平均収縮期血圧(SBP)と平均拡張期血圧(DBP)を、それぞれ求めた。

さらに、HRより心拍変動性(HRV)を算出し、特に心臓迷走神経活動のよい指標と考えられている高周波(HF)成分のパワー値比率を分析の対象とした。

3. 解析方法

PETによる脳画像は、SPM99を用いて標準的な手続きにより解析した。脳賦活部位の検出閾値は $p < .001$ (uncorrected)、クラスター閾値ボクセル数 20 以上に設定した。

被験者の意思決定・行動を検討するために、2つの刺激のうち、それぞれを選択した頻度を算出した。70%強化条件、50%強化条件それぞれ 120 試行のうち、有利刺激(ただし 70%強化条件において報酬確率の高い方を有利刺激とする)を選択した割合を、選択バイアスと呼び解析の対象とした。被験者個人ごとの選択バイアスについて、随伴性(70%、50%)×ブロック(1、2、3)の2要因繰り返しの分散分析(ANOVA)を行った。

脳波データより、報酬・罰の提示時点をトリガとした事象関連脳電位(event-related brain potentials)を解析した。原波形データに 20Hz の Low-pass フィルタをかけた後、トリガ・ポイントの前 400ms、後 1000ms の範囲について平均加算処理を行った。事象関連脳電位波形は、70%強化条件、50%強化条件のそれぞれにおいて、選択した刺激(有利、不利)、フィードバック(報酬、罰)について求めた。得られた成分について振幅を計測し、随伴性(70%、50%)×刺激(有利、不利)×フィードバック(報酬・罰)の3要因繰り返しの ANOVA を行った。

心臓血管系指標(HR、SBP、DBP、HRV)については、各ブロック前2分間の測定値の平均値を求めてベースラインとし、ブロック中2分間の測定値の平均値を課題中の反応とした。これらの値について、随伴性(70%、50%)×期間(ベースライン、課題中)×ブロック(1、2、3)の3要因繰り返しの分散分析(ANOVA)を行った。

脳活動と脳波、心臓血管系各指標の関連を検討するために、脳波については後述の事象関連脳電位成分の振幅を、心臓血管系指標について課題中または課題後の値からベースライン値を減じたものを変数とし、SPM99 を用いた回帰解析により相関脳画像を作成した。脳賦活部位の検出閾値は $p < .001$ (uncorrected)、クラスター閾値ボクセル数 20 以上に設定した。この閾値は相関解析としては非常に保守的なものであるが、多くの指標との相関画像を求めるために、タイプ1エラーの可能性を考慮して、そのように設定した。

C. 結果

1. 学習成績

図3に示すように、70%強化条件では、70%で強化される有利刺激を約 70%の割合で選択するようなバイアスがみられた。一方、50%強化条件では、両方の刺激が約 50%ずつ選択されていた。これは心理学でマッチング則(matching law)と呼ばれる行動の原理であり、意思決定は強化確率に対応して行われることが示唆された。ANOVA の結果、条件の主効果が有意であり($F(1, 15) = 9.28, p < .01$)、70%条件の方が有利刺激を選択する割合が高いことが示され、上記の視察の結果が確認された。

2. 脳血流量

70%強化条件と 50%強化条件において特異的に賦活した脳部位を同定するために、両条件から統制条件を減算する解析を行った。その結果、両条件に共通して、前頭眼窩野(BA11)、背側前部帯状回(BA24)及び内側前頭前野(BA32)、がそれぞれ有意に賦活した。しかし、それらの部位の賦活は、明らかに 50%強化条件においてより顕著であった。また、50%強化条件においては、淡蒼球の賦活も観測された(図4)。

両条件の賦活部位を詳細に比較すると、70%強化条件において賦活した左前頭眼窩野と内側前頭前野のピーク座標の位置は、50%強化条件のそれとほぼ完全に一致する。よって、賦活がみられた部位は随伴性によらず確率学習課題、特に意思決定過程を反映する神経ネットワークであり、随伴性が低い場合には、それが動員される程度が強いものと推測される。

3. 脳波

関心を持つ時間範囲の事象関連脳電位波形を視察すると、いずれの条件においても、Czにおいて、フィードバック刺激の提示後 400ms 付近に頂点を持つ陽性成分が得られた(図5)。これは P3と呼ばれる成分であると解釈される。P3 振幅についての ANOVA の結果、有意な効果は認められなかった。

しかし、P3成分の振幅と、PET により観測されたrCBFの相関画像を作成すると、70%強化条件においてのみ、前頭眼窩野、内側前頭前野・前部帯状回、淡蒼球などに広範な正の相関がみられた(図6)。この結果は、この課題中に観測された P3が、上記の前頭前野・基底核の神経ネットワークの活動を反映していることを示唆するものである。

4. 脳活動と課題成績の関連

我々は、これまで検討してきた脳活動や心臓血管系活動が、確率学習課題の成績とどのような関連を示すかに関心を持つ。

そこでまず、事象関連脳電位の P3 成分振幅について、次のように有利刺激と不利刺激における振幅

比を求めた。

振幅比 = (有利刺激の報酬/有利刺激の罰)/(不利刺激の報酬/不利刺激の罰)

この指標は、ふたつの刺激を選択した結果得られた報酬と罰を、どの程度弁別して評価していたかを表す。70%強化条件と50%強化条件のそれぞれにおけるP3振幅比と、課題における選択バイアスの間には、70%強化条件においてのみ、指数関数的な関連が見られた(図7: $R^2 = .64, p < .05$)。50%強化条件においては、なんらの関連もみられなかった(図7)。

これらの結果より、選択バイアスの大きい、すなわち有利刺激を多く選択して、その結果報酬を多く獲得できた個人は、P3に反映される刺激と報酬の随伴性評価が確立していたものと考えられる。

5. 心臓血管系指標

心臓血管系指標であるHR、SBP、DBPの、条件ごとの平均値と標準偏差を、表1に示す。HR、SBPにおいて有意あるいは有意傾向の、条件×期間の交互作用がみられた($F(1, 15) = 2.78, p < .10$; $F(1, 15) = 3.62, p < .05$)。LSD法により下位検定を行った結果、いずれの指標においても、ベースライン値に比べて課題中の値が増加していたことが明らかになった。さらに、その度合いは、70%強化条件において、50%強化条件よりも大きかった。すなわち、確率学習課題の急性ストレス負荷により、心臓血管系活動は亢進するが、その度合いは課題が学習不可能な場合には若干抑制されることが示唆された。

この現象の背後にある神経支配機構を推測するために、課題中のHRVのHF成分は50%強化条件で有意に大きいことが示された($F(1, 15) = 3.74, p < .05$)。これは、同条件において心臓迷走神経活動がより優勢であったことを示し、HR、SBP、DBPの相対的な亢進抑制は、迷走神経活動により媒介されていたことが示唆された。

6. 脳による心臓血管系の制御

脳は、自律神経活動の調整により末梢の心臓活動を制御していると考えられる。課題中における、そうした脳による心臓血管系の制御の様相を検討するために、PETにより観測されたrCBFと、HRVのHF成分の相関画像を作成した。

1) 70%強化条件

HF成分は、precuneusと小脳におけるrCBFと、有意な正の相関を示した(図8)。

2) 50%強化条件

HF成分は、precuneusに加えて、吻側前部帯状回(BA24)のrCBFと頑健な正の相関を示した(図8)。吻側前部帯状回は心臓迷走神経の支配部位として知られており、fMRIを用いた神経イメージング研究

でもHRVのHF成分とこの部位の賦活の関連が報告されている(Critchley, 2005)。よって、随伴性の学習が困難で、意思決定のために継続的な対処が要求される50%強化条件においては、吻側前部帯状回由来で心臓迷走神経活動が亢進され、その結果、課題中のHRやBPの亢進を、ある程度抑えるように調整がはかれることが明らかになった。

D. 考察

1. 急性ストレス課題としての確率学習課題

本研究で使用した確率学習課題は、心拍、血圧を頑健に上昇させ、心臓血管系に典型的な急性ストレス反応を惹起したことが明らかになった。これは、同様な課題を用いた我々の先行研究の結果とも一致している(Kimura et al, in press.)。

本研究では各ブロックにおける負荷は4分と、比較的短時間であったが、実用上十分なストレス反応を惹起しうることが示された。これは、本研究で用いられた確率学習課題が比較的単純な2刺激間の選択課題でありながら、金銭的インセンティブを伴う一種のギャンブルとして設定されていたため、及び700msという時間圧のもとで選択を迫るものであったためだと推測される。こうした点を適切に設定・操作すれば、この課題は、今後ストレス負荷の評価システムを開発する上で有効な手段になりうると考えられる。

PETを用いた神経イメージング研究においては、多くの実験要因は被験者内要因として操作される。その場合、条件をカウンタ・バランスするなどして統制することが普通である。本研究でも、70%強化条件と、50%強化条件という2種類の実験条件を設定したが、これは被験者内条件として操作し、順序にカウンタ・バランスを導入した。その結果、いずれの指標にも順序効果はみられず、この課題は被験者内要因の操作が容易であることが示唆された。これは、各々の条件の直前に、それに対応した統制条件を挿入することで、前の条件の影響がキャンセルされたものと考えられる。この点でも、本研究は神経イメージング研究に適したものであるといえよう。

2. 刺激－報酬の随伴性と脳活動

行動的結果をみると、選択バイアスは70%強化条件では平均70%、50%強化条件では平均50%であり、平均値をみるかぎり、いわゆるマッチング則にそう反応傾向がみられた。この行動傾向は、特に70%強化条件では経済学的な意味で「合理的」ではない。なぜならば、有利刺激を100%選択することが、もつとも報酬獲得の期待値が高いからである(70%)。マッチング的選択は、期待値を58%に減らしてしまう。さらに、図7にみられるように、選択バイアスにはかなりの個人差がみられる。すなわち、上記の「合理的」な選択パターンが顕著な個人から、逆にむしろ不利刺激の方を多く選択している個人さえ存在する。この個

人差は、学習の能力、刺激－報酬の随伴性を評価した上での意思決定方略など、さまざまな要因を反映していると考えられる。今後、それらを詳細に検討することによって、本研究で使用した確率学習課題を慢性ストレスが認知機能に及ぼす影響を評価するツールとして開発していくことが可能であると考えられる。

さらにこのとき、そうした個人の意思決定は、P3に代表される事象関連脳電位の成分に鋭敏に反映されることが、本研究により初めて明らかになった。各刺激を選択した後、それがもたらす報酬・罰のフィードバックに同期して観測される P3 成分は、それぞれの刺激選択が、どの程度報酬や罰をもたらすかという予期と、実際の結果との相対的な関係を反映すると考えることができよう。すなわち、ひとつの刺激の報酬への P3 成分と罰への P3 成分の振幅比は、その刺激への評価を反映しており、その振幅比の刺激ごとの比は、「一方の刺激が他方の刺激に比べて、どれくらいよさそうか」という評価を反映している。図8A にみられるように、この P3 振幅比と選択バイアスは極めて強い関連がある。また事実上学習が不可能な 50%強化条件では、なんらの関連もみられていない。これらの結果は、P3を手がかりにして脳内での意思決定機構を検討していく可能性を強く示唆する。

さらに指標としての P3の妥当性は、70%強化条件においてのみ、前頭眼窩野、内側前頭前野・前部帯状回、基底核における血流量と P3振幅が有意な相関を示したことから支持される。これらの脳部位は、PET 画像の減産解析では、逆に 50%強化条件において有意で頑健な賦活がみられた所である(図4)。これは、50%強化条件においては、被験者は学習が不可能であったために最後まで探索を続けていたために持続的な脳賦活が生じていたためだと考えられる。70%強化条件では随伴性の評価が確立すれば、上記の前頭前野－基底核のネットワーク賦活は低下する。それとともに P3 振幅も低下したために、結果として正の相関が出現したと考えることができるだろう。すなわち、個人の P3 振幅は、前頭前野－基底核ネットワークによる刺激－報酬の随伴性評価の過程を表しているといえる。

前頭眼窩野はすべての感覚入力 of 最終的な投射を受け、その報酬値を決定する機能がある。内側前頭前野は、行動の適応性をモニターし、エラーを修正する機能がある。これらの領域は双方向に密接な神経連絡を有し、ひとつのネットワークとして機能するといわれている(Kringelbach, 2005)。また淡蒼球や線条体などの基底核は、報酬予期やその誤差の計算、それに基づく反応選択のバイアスなどに関連することが報告されている(O'Doherty, JP, Dolan, RJ)。本研究で観測された脳賦活は、こうした神経ネットワークの活動をとらえたものと考えられる。

3. 刺激－報酬の随伴性と心臓血管系活動

50%強化条件では、心臓血管系の急性ストレス反応が、いくぶん抑制される方向に調整されていた。この調整のメカニズムとしては、交感神経活動の減衰、副交感神経活動の亢進、及びその両方が考えられる。しかしここでは、副交感神経活動の方が、はるかに時間的反応性が高いこと、副交感神経活動は前頭前野領域により直接支配され、自律系活動を常時抑制的に制御していること(Thayer & Brosschot, 2005)、前述したように本研究においてこの条件で前頭前野の賦活が顕著であったことから考えて、副交感神経系の関与が大きいことが推測される。心臓迷走(副交感)神経系活動を反映する指標である HRV の HF 成分のパワー値比が、この条件で有意に大きかったことも、この推測を支持している。

さらに、50%強化条件のみにおいて、HRV の HF 成分が吻側前部帯状回における脳血流量と相関を示したことも、上記の推測と整合する(Critchley, 2005)。この脳部位は、視床下部、中脳水道灰白質、中脳の各神経核に直接神経投射し、それらの機能を調整する機能がある(Kringelbach, 2005)。さらに、この脳部位は、前頭眼窩野とも密な神経ネットワークを形成している。これらを勘案すると、本研究の結果は、刺激－報酬の随伴性の評価に基づいて、心臓迷走神経系を介して、オンラインで末梢の心臓血管系反応を調整する機能を初めて画像化して示したものであるといえる。

これらの脳領域が適切に働くことが、常時変化する環境に柔軟に適応する上できわめて重要であると考えられる。実際、ストレス事態での前頭前野による副交感神経系制御の機能低下は、多くの身体的・精神的疾患につながるリスク要因であることが指摘されている(Thayer & Brosschot, 2005)。このことから、急性ストレス負荷事態において、刺激－報酬の随伴性を逆転させたり、段階的に変化させるような操作を導入し、それへの行動、脳賦活、心臓血管系の反応性の鋭敏さにより、慢性ストレス負荷の程度を推測することが可能かもしれない。次年度では、過重労働により慢性ストレス状態にある個人と、そうでない個人について、そうした比較を行うことにより、この仮説を直接検証することを試みる。

E. 結論

PET、脳波、心臓血管系活動の同時計測により、確率学習課題が急性ストレス負荷として有効なこと、報酬・罰の随伴性に応じて、脳は柔軟にそれを評価し、心臓血管系活動とあいまって学習成績に影響することが示された。この手法をより洗練し、基礎的知見を蓄積することで、過重労働などの、慢性ストレス負荷の評価システムの開発につながる可能性がある。

F. 健康危険情報

該当せず。

G. 研究発表

1. 論文発表

Isowa T, Ohira H, Murashima M. Immune, endocrine and cardiovascular responses to controllable and uncontrollable acute stress. *Biol Psychol.* 2006; 71: 202-213. Isowa T 2006

Ohira H, Nomura M, Ichikawa N, Isowa T, Iidaka T, Sato A, Fukuyama S, Nakajima T, Yamada J. Association of neural and physiological responses during voluntary emotion suppression. *Neuroimage.* 2006; 29: 721-733.

2. 学会発表

Ohira H, Ichikawa N, Isowa T, Nomura M, Kimura K, Kanayama N, Fukuyama S, Shinoda J, Yamada J. To learn is somatic: brain-body association during stochastic learning. 12th Annual Meeting of Human Brain Mapping. (Florence, Italy), July 2006.

Ohira H, Ichikawa N, Isowa T, Nomura M, Kimura K, Kanayama N, Fukuyama S, Shinoda J, Yamada J. Brain Activity during Reward Learning: simultaneous Recording of PET and peripheral physiological activity. 46th Annual Meeting of Society for Psychophysiological Research. (Vancouver, Canada), November 2006.

H. 知的財産権の出願・登録状況(予定を含む)
該当せず。

I. 引用文献

Critchley HD. Neural mechanisms of autonomic, affective, and cognitive integration. *J Comp*

Neurol. 2005; 493: 154-166.

Isowa T, Ohira H, Murashima M. Immune, endocrine and cardiovascular responses to controllable and uncontrollable acute stress. *Biol Psychol.* 2006; 71: 202-213.

Kimura K, Ohira H., Isowa T, Matsunaga M, Murashima S. Regulation of lymphocytes redistribution via autonomic nervous activity during stochastic learning. *Brain, Behav, Immun.* in press.

Kringelbach ML. The human orbitofrontal cortex: linking reward to hedonic experience. *Nat Rev Neurosci.* 2005; 6: 691-702.

O'Doherty JP, Dolan RJ. The role of human orbitofrontal cortex in reward prediction and behavioral choice: insights from neuroimaging. In Zald, DH, Rauch, SL. (Eds.), *The orbitofrontal cortex.* New York: Oxford University Press. 2006, pp. 265-284.

Ohira H, Nomura M, Ichikawa N, Isowa T, Iidaka T, Sato A, Fukuyama S, Nakajima T, Yamada J. Association of neural and physiological responses during voluntary emotion suppression. *Neuroimage.* 2006; 29: 721-733.

Thayer JF, Brosschot JF. Psychosomatics and psychopathology: looking up and down from the brain. *Psychoneuroendocrinology.* 2005; 30: 1050-1058.

表 1 心臓血管系指標の平均値と標準偏差

	70%強化条件		50%強化条件	
	ベースライン	課題中	ベースライン	課題中
HR (bpm)	65.38 (7.99)	66.83 (9.55)	63.75 (7.99)	64.85 (8.27)
SBP (mmHg)	129.71 (11.08)	133.62 (6.75)	127.02 (10.41)	129.54 (10.77)
DBP (mmHg)	62.02 (6.91)	64.20 (6.75)	59.36 (8.41)	62.34 (7.79)
HRV HF 成分 (%)	47.93(9.19)	53.09(9.71)	51.79(10.51)	58.19(11.99)

括弧内は標準偏差

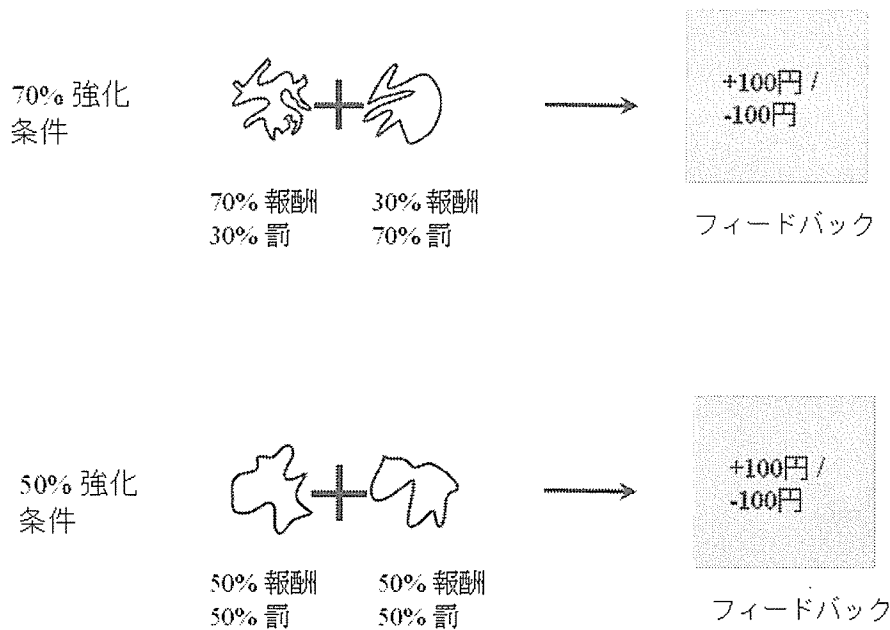
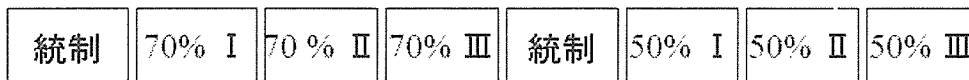


図1 本研究で使用した確率学習課題：被験者は、2つの無意味な図形のうち、一方を選択するよう求められる。70%強化条件では、一方の刺激選択は70%の金銭的報酬(100円)、30%の金銭的罰(-100円)をもたらす。50%強化条件では、どちらの刺激選択も、50%の金銭的報酬と、50%の金銭的罰をもたらす。

1. 初期70%強化群(N=8)



2. 初期50%強化群(N=8)

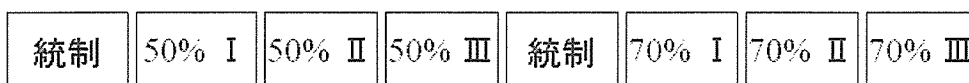


図2 本研究の実験デザイン：実験は、8ブロックから構成されていた。70%強化条件、50%強化条件は、それぞれ連続的に3ブロック、合計120試行が遂行された。両条件の順序は被験者間でカウンタ・バランスされた。また、両条件の直前には、コンピュータが選択刺激を指定する統制条件が置かれた。統制条件における強化確率は、それに続く3ブロックと同一とされた。

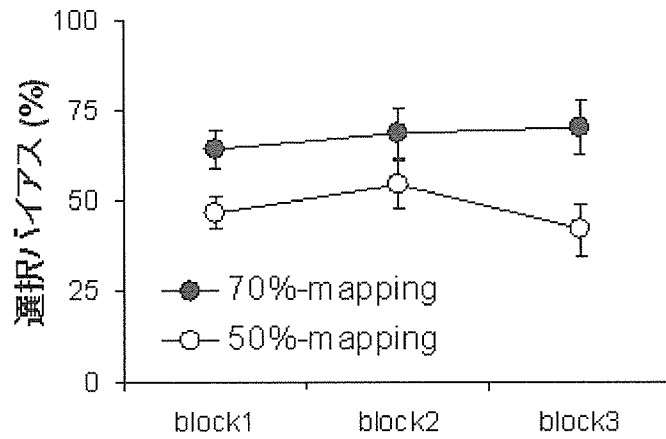


図3 確率学習課題における選択バイアス：70%強化条件においては、有利刺激(70%の確率で報酬をもたらす刺激)の選択率を示す。50%強化条件においては、70%強化条件と対応した刺激の選択率を示す。

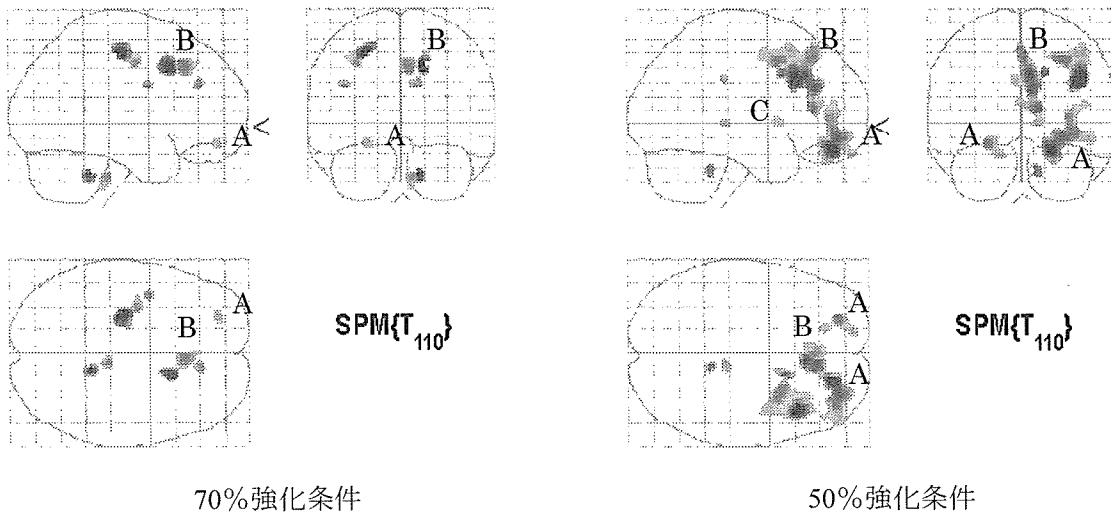


図4 70%強化条件と50%強化条件における脳賦活 ($p < .001$, uncorrected)：A：前頭眼窩野、B：内側前頭前や、C：淡蒼球。

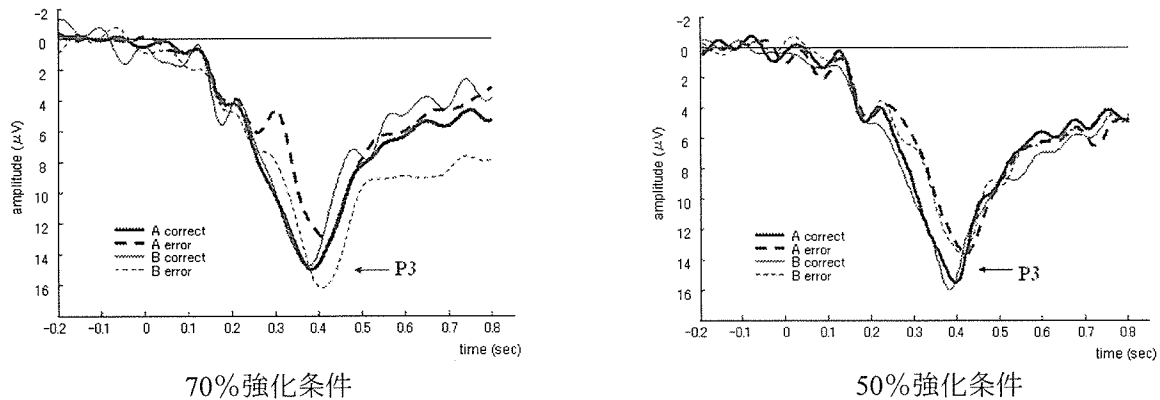


図5 報酬・罰のフィードバックに同期した事象関連脳電位(Cz)：両条件において、明瞭なP3と思われる陽性成分が観測されている。

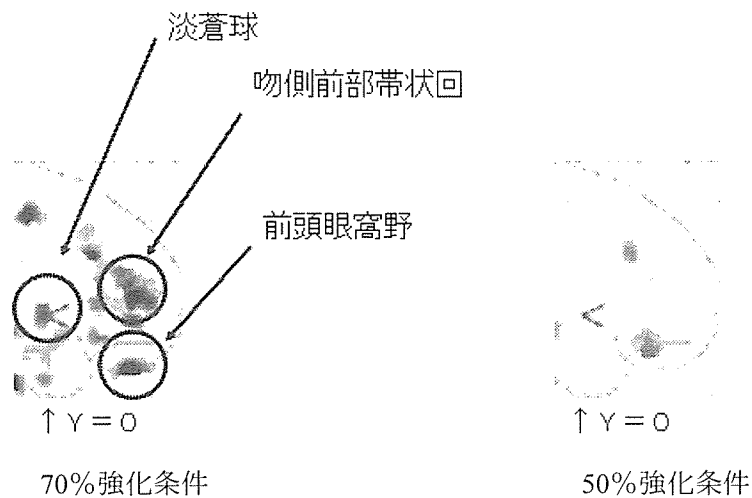


図6 P3振幅と、脳血流量の相関画像：70%条件において、顕著な正の相関が観測された。

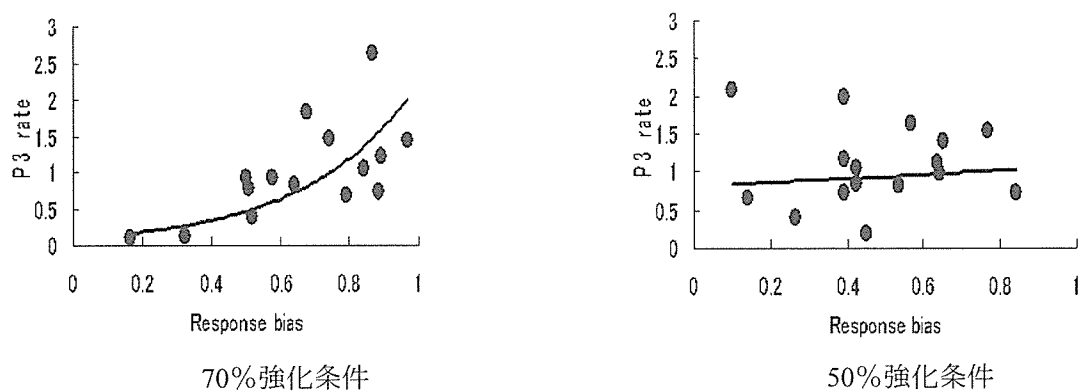


図7 P3振幅比と選択バイアスの関連：ただし、P3振幅比 = (有利刺激への報酬/有利刺激への罰) / (不利刺激への報酬/不利刺激への罰)、選択バイアス = 有利刺激を選択した割合。

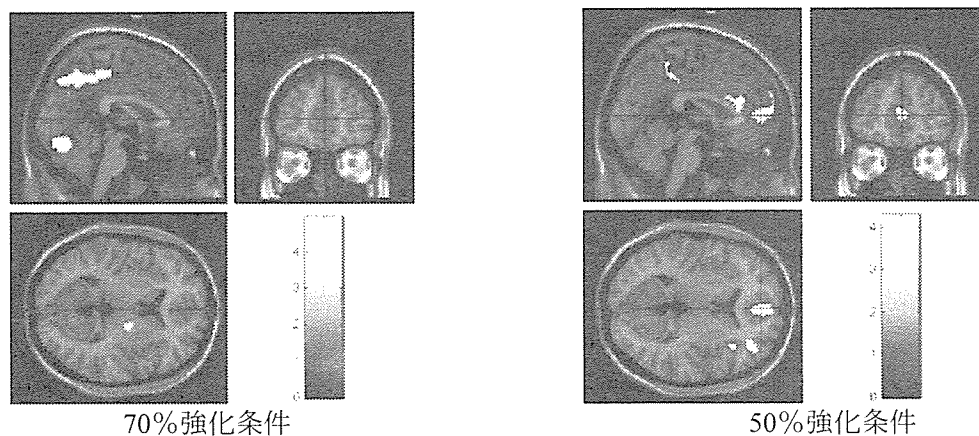


図8 心臓迷走神経活動を反映する心拍変動性の高周波成分のパワー値と脳血流量の相関画像。

平成 18 年度厚生労働科学研究労働安全衛生総合研究事業
「過重労働等による労働者のストレス負荷の評価に関する研究」
分担研究報告書

過重労働等によるストレス負荷評価のための新しい調査票の開発に関する研究

分担研究者 岩田 昇（広島国際大学心理科学部・教授）
研究協力者 菊地賢一（東邦大学理学部・助教授）
藤原裕弥（東亜大学総合人間文化学部・講師）

研究要旨

平成 18 年度当該研究事業の分担研究として、①平成 17 年度の研究課題で開発した項目反応理論に基づく適応型テストあるいは Computerized-Adaptive Testing (CAT) の改訂版の追加調査、ならびに②うつ症状測定用の CAT の開発およびその信頼性・妥当性検討を行った。

①研究 1 では、前年度のコンピュータ型の簡易システムに身体症状の測定モジュールを追加した改訂版を開発した。これを昨年度の事業所とは異なる 1 事業所で試用し、50 名の回答データを解析した。その結果、ストレス反応の CAT 測定は昨年度と同様の結果が得られ、ストレスレベルは過重労働状況と有意に関連していた。また、新たに導入した身体症状モジュールでも、ストレスレベルや過重労働状況との有意な関連が見られ、身体症状モジュールの意義が確認された。

②研究 2 では、ストレス反応測定評価のための CAT モジュールの検討のために、抑うつ尺度および追加項目（計 24 項目）の CAT システムを構築し、大学生 209 名に試用した。うち 168 名には約 1 週間の間隔で再度回答を求めた。また、同一画面形式で全項目に回答を求める CBT も同時に実施した。その結果、うつ測定 CAT の十分な再テスト信頼性と高い並存的妥当性が確認された。これにより、特にうつ症状に焦点を当てたストレス反応測定モジュールが整備され、次年度の総合システムの構築に向けた準備がほぼ整ったと考えられた。

【研究 1】過重労働等によるストレス評価用コンピュータ型簡易システム開発に関する研究

A. 研究目的

従来の職業性ストレスや精神健康調査票の測定方法が依拠している「項目固定型テスト」を前提とした古典的テスト理論 (Classical Test Theory、以下 CTT) には、1) 尺度単位でしか測定評価ができない、2) 信頼性・妥当性も尺度全体 (得点) でしか保障されないなどの制約があった。そのため、例えば職域メンタルヘルス活動などで同一尺度を繰り返し施行した場合、その心理測定法的特性が低下することが懸念されても、尺度得点の互換性の制約のために別の尺度を使うことはできず (CTT では比較不可能)、改良すべき課題となっていた。

このような制約を受けない柔軟なストレス評価を可能にする方法論が、項目反応理論 (Item Response Theory、以下 IRT) である。IRT では、測定しようとする潜在特性に対する個々の項目の反応特性を明らかにし、同時にその潜在特性上での各回答者のレベル (θ) を推定することができる。項目単位で測定評価ができるという IRT の特徴は、一次元性が確認された項目群であれば、別々の尺度の結果を対比をも可能にする。IRT を適用したテスト構成の例は、TOEFL などで見られる。

「適応型テスト」 (Computerized-Adaptive Testing、以下 CAT) は、IRT を応用した新世

代の心理測定法である。CATでは、逐次提示される項目への回答に基づき、その時点での θ を推定し、事後分散を最小にする項目が次に選択され提示される。この手続きを繰り返して、 θ の推定誤差が予め設定した範囲内に達した時点で項目提示が終了する。これがIRTに基づくCATの基本的なシステムである。CATでは、必ずしも回答者全員が同じ項目に回答するとは限らず、推定終了までの項目は少ないため、回答者の負担は軽減される。

このようにCATには多くの利点があるが、用いることができる項目はあくまでも次元性が認められた項目群に限られるという制約条件も存在する。過重労働状況、職業に伴うストレスナーなどは、必ずしもすべて一次元的な測定概念ではない。また、CAT化しなくとも現場での簡便な把握のためには、従来型の測定尺度でも有用な場合もある。

そこで平成17年度の分担研究課題として、労働者向けの測定ツールとして、ストレス反応の測定にCATを採用し、過重労働状況、ソーシャルサポートの測定にCBT(Computer-Based Test: 通常の質問項目を固定した順に提示する従来型の測定法のコンピュータ版)を採用した簡易システムの試作品を開発した。さらに、現場試用として、A事業所36名を対象にデータ収集ならびに解析を行なった。しかし、システムの設定条件に誤りが発見されたため、今年度当該部分を修正し、新たにB事業所50名で試用した。

B. 研究方法

1. CATの現場における回答データ

九州地区のB社に協力を依頼し、2006年5月下旬よりシステム改訂版を試用した。B社の健康診断の実施会場の待合エリアに、特設テーブルを設置し、デスクトップ・コンピュータ1台を置いて試用協力を求めた。その結果、56

名の回答が得られた。

ストレス反応項目への回答時間を検討したところ、そのうち6名が平均1秒未満で回答していた。平成17年度の解析でも、項目あたり平均応答時間1秒未満のデータは除外しており、この6名を項目内容の理解が不十分な不適切回答者と判断し、残り50名(男性28、女性22)の回答データを解析した。対象者の平均年齢(±SD)は、男性49.6±17.7歳、女性35.7±16.0歳であった($p<.01$)。

(倫理面への配慮)

事業所での試用に際して、性・年齢(年齢の詳細を明らかにしたくない場合には、99と入力)の情報のみ、フェイスシート画面に提示し、同画面上部には、この測定は強制的調査ではないこと、協力可能な方のみ回答を求めたい旨を明記した。

C. 研究結果

1. 現場試用結果(ストレス測定CAT)

今回試用したCATに基づく回答者の平均ストレスレベル(θ 値=ここでは偏差値換算)は54.7±12.6で、40未満の者が2名、40~60未満が37名、60を超える回答者が11名と、全体としてはやや高いストレスレベルにあった。

表1にCAT終了条件とした θ の推定誤差0.35という収束条件までの所要回答項目数の分布を示す。平均項目数は10.7±3.5項目で、約8割の回答者が12項目までで収束していた。測定に要する時間は平均65.2±36.6秒で、女性(54.4±21.2)よりも男性(73.7±43.7)の方が有意に長かった($p<.05$)。測定に使用された頻度を見ると、明らかに識別力の高い項目ほど頻繁に採用されていた(表2)。

2. ストレス関連要因

表3にCAT/CBTシステムの評価指標と各測

定項目の回答および回答時間との相関係数を示す。ストレスレベル（ θ 値）は年齢との関連を認めなかったが、過重労働状況と正の、ソーシャルサポートとは負の相関が見られた。

個々の項目で見ると、CATの前に固定して配置した3逆転項目のうち、#3「生き生きしている」がストレスレベルとの相関を認めなかったが、その他の提示された項目はすべて有意な相関を示していた。CBTとしてCATの後に配置した身体症状では、11項目中8項目がストレスレベルと有意な相関を示し、特に#1「めまい」、#3「頭重・頭痛」、#8「胃腸」、#9「食欲」、#11「不眠」が高い相関を示した。

過重労働状況8項目では、1項目が有意傾向の相関だった他は、すべてストレスレベルと有意な相関を示し、特に#5「仕事のことで悩む」、#8「家でゆっくりくつろぐ（逆転項目）」の相関が高かった。ソーシャルサポート6項目では、#1「職場の上司・同僚（含、部下）に気軽に話ができる」、#2「配偶者や家族・友人に気軽に話ができる」の2項目のみストレスレベルと有意な相関を示した。なお、これらのモジュールの項目の中で、項目回答時間がストレスレベルと有意な相関を示したものはなかった。

3. 過重労働状況と関連するストレス症状

過重労働状況と有意な相関を示したストレス症状項目は、CATの方では#12「落ち着かない」、#14「面倒だ」、#15「集中できない」の3項目で、身体症状では#1「めまい」、#2「体の節々が痛む」、#4「首筋や肩こり」、#8「胃腸」、#10「便秘や下痢」であった。これらから、過重労働状況は精神的なストレス反応よりもむしろ身体的な反応の方が関連していることがうかがえる。

ただ、前者はCATのため、N数が少ない項目も散見されるのに対し、後者はこの改良版で試みにCBTの別モジュールとして組み込まれ

ているため全回答者のデータが得られている。相関係数の大きさだけを見ると、CATの方にもいくつかN数のために有意水準に達していない可能性がある項目も推察される。さらにサンプル数を増やして検討する必要がある。

D. 考察

ストレス反応測定のためのIRT-CATシステムを改良し、身体的ストレス反応を測定評価するためのCBTモジュールを組み込んだ改訂版を労働者50名に試用した結果、推定誤差0.35未満の収束設定では、Pretestの3項目を除けば、6~7項目程度で終了していた。これは昨年度の改訂前システムの試用結果とほぼ同様であり、簡便な測定ツールとしてのCATの利点が再確認された。

識別力の低い、いわゆる傾きの緩い項目である「身体愁訴」項目群は、推定誤差の収束をCATアルゴリズムでは選択されないが、CATで推定したストレスレベルとは非常に高い関連を示した。このことより、昨年度の考察した身体症状の別モジュール構築の意義は支持されたと考えられる。また、過重労働状況との関連性を見ても、身体症状を積極的に測定する意義は明白である。

一方、過重労働状況がCAT推定のストレスレベルと関連していることが明らかとなった。ただし、今回の解析では全体の平均的関わり程度としての相関（線形回帰）でしか検討しておらず、特に高ストレスレベルに影響する過重労働状況項目の検討や、その項目のどのレベルから問題になるのかなど、今後より詳細に解析していく必要がある。

なお、各項目の回答反応時間について、ストレスレベルその他との関連性を相関分析によって探索したが、関連する項目は見出されなかった。コンピュータ仕様のこの種のテストの大きな特徴の一つとして、この回答反応に関わる

時間変数が挙げられるが、従来のストレス研究の枠組を超えたこの評価軸をストレス測定評価の軸に組み込むことの積極的な証拠は未だ得られていない。

E. 結論

ストレス反応、身体愁訴、および過重労働状況、ソーシャルサポートの4側面を同時にかつ簡便に測定評価することができるコンピュータ型の簡易システムの試作品を開発した。ストレス反応に関するモジュールは、適応型テスト (IRT-CAT) システムで、それ以外のモジュールはCBT形式で設定した。また最後に相談希望の質問と、回答結果のフィードバック画面を加え、システム初版を完成させた。

1事業所50名の試用協力を得、回答データを解析した。ストレス反応のCAT測定は昨年度と同様の結果が得られ、ストレスレベルは過重労働状況と有意に関連していた。また、新たに導入した身体症状モジュールでも、ストレスレベルや過重労働状況との有意な関連が見られ、身体症状モジュールの意義が確認された。

労働者個人を対象としたコンピュータ仕様の測定ツールの試作システムが構築された。今後、本複合測定システムを現場試用し、検討していく予定である。

F. 健康危険情報

なし。

G. 研究発表

1. 論文発表

なし。

2. 学会発表

Iwata N. Measurement in East and West: how differ does the Japanese response to the assessment tool? The 26th International Congress of Applied Psychology, Athens, July 2006.

H. 知的財産権の出願・登録状況

なし。

表1 ストレス反応測定CATの収束までの提示項目数

A社 (Mar 2006)				B社 (May 2006)			
提示項目数	収束回数	%	累積%	提示項目数	収束回数	%	累積%
7	1	2.9	2.9	7	2	4.0	4.0
8	13	38.2	41.2	8	11	22.0	26.0
9	8	23.5	64.7	9	8	16.0	42.0
10	5	14.7	79.4	10	9	18.0	60.0
11	1	2.9	82.4	11	6	12.0	72.0
—	—	—	—	12	4	8.0	80.0
13	4	11.8	94.1	13	5	10.0	90.0
—	—	—	—	15	3	6.0	96.0
—	—	—	—	16	1	2.0	98.0
18	2	5.9	100.0	—	—	—	—
—	—	—	—	29	1	2.0	100.0
合計	34	100		合計	50	100	

平均所要時間(秒) 65.2
SD 36.6

表2 ストレス反応測定CATにおける各項目の採用頻度

項目	内容	閾値			識別力 a	2006Mar	2006May	備考
		b1	b2	b3		A社 頻度 (%)	B社 頻度 (%)	
BSR1	活気がわく	-3.49	-0.99	1.98	0.54	34 (100)	50 (100)	preset
BSR2	元気いっぱい	-2.63	-0.63	1.84	0.63	34 (100)	50 (100)	preset
BSR3	生き生き	-3.02	-0.90	1.58	0.63	34 (100)	50 (100)	preset
BSR4	怒り	-1.31	0.78	2.80	0.01	0 (0)	5 (10)	a=0.79
BSR5	腹立	-1.06	0.77	2.45	0.95	5 (15)	10 (20)	
BSR6	イライラ	-1.00	0.64	2.13	1.15	29 (85)	43 (86)	
BSR7	ひどく疲れた	-1.44	0.36	1.82	1.01	7 (21)	12 (24)	
BSR8	へとへと	-0.24	1.08	2.31	0.99	4 (12)	9 (18)	
BSR9	だるい	-0.91	0.67	1.98	1.12	23 (68)	34 (68)	
BSR10	気が張りつめている	-2.06	0.65	3.15	0.48	1 (3)	1 (2)	
BSR11	不安だ	-0.43	0.98	2.13	1.15	34 (100)	47 (94)	
BSR12	落ち着かない	-0.14	1.23	2.42	1.30	34 (100)	50 (100)	
BSR13	憂うつだ	-0.28	0.95	1.93	1.65	34 (100)	50 (100)	
BSR14	面倒だ	-0.16	1.56	2.79	1.13	11 (32)	26 (52)	
BSR15	集中できない	-0.37	1.67	3.04	1.07	7 (21)	14 (28)	
BSR16	気分が晴れない	-0.66	0.99	2.01	1.59	34 (100)	50 (100)	
BSR17	仕事が手につかない	0.45	2.31	3.74	0.99	1 (3)	5 (10)	
BSR18	悲しい	0.67	2.07	3.26	0.96	1 (3)	5 (10)	
BSR19	めまい	1.38	3.13	4.80	0.69	0 (0)	13 (26)	
BSR20	体の節々が痛む	1.14	3.42	5.22	0.49	0 (0)	1 (2)	
BSR21	頭重・頭痛	0.29	2.30	4.19	0.65	2 (6)	3 (6)	
BSR22	首筋や肩こり	-1.09	0.92	2.57	0.48	1 (3)	1 (2)	
BSR23	腰痛	-0.18	2.10	4.13	0.40	0 (0)	1 (2)	
BSR24	眼の疲れ	-1.83	0.43	2.39	0.52	1 (3)	1 (2)	
BSR25	動機息切れ	1.58	3.65	5.33	0.60	0 (0)	3 (6)	
BSR26	胃腸	0.31	2.23	3.84	0.60	0 (0)	1 (2)	
BSR27	食欲	1.11	3.18	4.70	0.65	0 (0)	3 (6)	
BSR28	便秘や下痢	0.23	2.48	4.49	0.46	0 (0)	1 (2)	
BSR29	不眠	0.31	2.17	3.75	0.65	1 (3)	3 (6)	

【研究2】うつ症状測定用のCATの開発およびその信頼性・妥当性検討

A. 研究目的

MEDLINE・PsychLit・PsychInfo を用いて医学・保健領域におけるCATの開発・適用例を検索してみると、2007年3月12日現在まで、当該研究報告はわずか7編しか見られない。そのうち、メンタルヘルス測定に用いているものは2編 (Fliege ら, 2005 ; Walter ら, 2007) だけで、いずれもベルリン大学医学部の同じ研究グループからの報告である。また、残りの5編はリハビリテーション領域の機能検査測定に応用したものである。

このように、メンタルヘルスに関連したCATの研究は、まだほんの緒についたばかりである。そこで本研究2では、うつ症状測定を目的としたCATの開発および信頼性・妥当性の検討を試みた。これは同時に、過重労働に関連するストレス負荷状況に関する簡易測定システム (上記研究1) の中核をなす精神的ストレス反応測定モジュールの検討でもある。

B. 研究方法

1. CAT開発のための基礎データ

総計2,142名 (男1,682、女460) の労働者および大学生の回答データをIRT分析に用いた。平均年齢は男性37.4 (SD 11.7) 歳、女性28.8 (SD 10.0) 歳であった。労働者は定期健康診断の際に、他のストレス尺度などを含む自記式質問紙調査票に回答した。大学生の方は、講義後に無記名で回答を求めた。

2. うつ症状尺度

Center for Epidemiologic Studies Depression Scale (CES-D: Radloff, 1977) を用いた。CES-Dは、過去1週間に経験した抑うつ症状20項目の頻度を測定するであり、16のネガティブ項目 (例、「憂(ゆう)うつだと感じる」) と4ポジテ

ィブ感情項目 (例、「毎日が楽しい」) より構成されている。ネガティブ項目は各7項目の『うつ感情』、『身体症状』、2項目の『対人関係』の3下位尺度に要約できる。回答者は各項目に対して、1週間(7日間)のうち、「めったにまたは全くない(1日未満)」、「いくらかまたは少しはある(1-2日)」、「ときどきまたはかなりある(3-4日)」、「たいていまたはいつもある(5日以上)」の中から1つ選択する。各回答は順に0~3と配点される。4ポジティブ項目ではこれらを逆転配点し、ネガティブ項目と同様に、スコアが高いほど抑うつ症状が高いようにする。CES-Dの標準的な区分点は、合計スコア ≥ 16 であり、これを満たすとうつの「高リスク群」とされる (Radloff, 1977)。

本研究では、日本語版CES-D (島ら, 1985) を用いたが、日本人におけるポジティブ項目の心理測定法的特性の問題が異文化間比較 (Iwata ら, 1994・1995 ; Iwata & Higuchi, 2000 ; Iwata & Buka, 2002) や臨床患者 (Iwata ら, 1998) で明らかになっていることから、4ポジティブ項目をネガティブ表現に改訂した項目 (Iwata ら, 1998) も付加した。本稿では、従来のCES-Dと区別し、CES-DRと記すことにする。

3. IRT解析

まずIRTの最も基本的な仮定である「次元性」を因子分析によって確認した。次いで、多値型のIRTモデル (Modified Graded Response Model、以下MGRM、Muraki, 1992) を適用し、選択肢間の閾値(位置パラメータ)および識別力を推定した。

MGRMモデルでは、4選択肢に対する回答確率をx軸(潜在特性)上の θ 値に対応させて推定する。MGRMでは、この4選択肢の一番左(軽症状レベル: 1) 対その他右側(選択肢

2～4)の境界を表す曲線、選択肢を真ん中で2分割する、軽症状(選択肢1・2)対その他(選択肢3・4)の境界を表す曲線、左3選択肢(1～3)対一番右(重症レベル:4)の境界を表す曲線を求める。これらから、各項目の各選択肢の θ 上の位置を求める。この交点 $b_1\sim b_3$ を項目の閾値とし、曲線の立ち上がりに対応する傾き(識別力)と共に項目の特性を示すパラメータとなる。

4. CATにおける θ 推定方法

CATでは回答データが得られるごとに、回答者の θ 値を逐次的に推定する。つまり、回答ごとに回答者の潜在特性上の値が変化していく。例えば、抑うつレベルという潜在特性を測定評価するCATの場合、「ゆううつだ」という項目に「ほとんどなかった」と回答すれば、それまでの抑うつレベル直線上の回答者の位置は、それ以前の位置よりもやや左(すなわち低レベル)側に移行し、逆に「ほとんどいつも」と回答すれば、それまでの抑うつレベル直線上の回答者の位置は、それよりもやや右(すなわち高レベル)側に移行していく。

今回のCATシステムでは、「ベイズ法」を用いた。逐次回答データから求めた θ の事後(確率)分布の平均値を θ 推定値とし、その事後分布の標準偏差を θ の推定誤差とするものである。

5. CATにおける項目選択方法

CATでは回答者の θ 推定値を受けて、次の質問項目を選択する。この項目選択には、いくつかの方法があるが、ここでは「制約付きベイズ法」を用いた。これは、ある θ 値がそこまでの回答で推定されているとして、事後分散が最も小さくなるという基準で次の項目を選択する方法である。この選択を行なうことにより、 θ 値の推定誤差は徐々に小さくなり、信頼に足

る推定値まで項目を提示していく。本研究ではCATの収束条件として、 θ の推定誤差0.30を採用した。これはMGRM解析での θ 推定値の標準誤差0.29に準拠し、設定したものである。

6. 本研究に用いたコンピュータ測定システム

図1に本研究のコンピュータ測定の手続きおよび抽出データを示す。今回はCATの前に、CES-DRのCBTを配置した。これはCATに基づく推定値 θ_{CAT} と全項目を用いるCBTに基づく推定値 θ_{CBT} との一致程度によって、CATシステムの測定精度を検討するための手続きである。なお、CES-DRのCBTの最初に2項目(「退屈だと思う」「ムカツク」)を練習課題として配置した。これらはその後のCBT・CATと同一の画面表示形式で、実験者がその画面を見せながら回答手順について説明した。

7. CAT・CBTシステムを用いたデータ収集

広島県および山口県下にある2大学において、大学生に本システムへの回答協力を依頼した。1週間程度の間隔において2度測定するという条件に協力同意が得られた209名(男101、女108)を対象に、本システムを試用した。この試用では、実験室に回答協力者および実験者の各1名のみでの状況下で、PCI台を用いて行なった。2回目の測定は、協力同意は得られていたものの、翌週の協力が急に不可能となった・予定日時を忘却したなどの理由により、168名(男80、女88)であった。平均年齢(±SD)は 21.2 ± 1.8 歳であった。再テストまでの間隔は、平均 6.6 ± 1.9 日であった。

なお、MGRMのパラメータに基づいて推定する θ_{CAT} 値および θ_{CBT} 値以外に、CBTへの回答から通常の集計方法によるCES-Dおよびポジティブ4項目の替わりに、そのネガティブ表現改訂項目を用いた場合の合計点も算出し、これらのスコア間の関連性を検討した。

(倫理面への配慮)

本研究は、既存データの再解析から得られたパラメータ値を用いたシステム開発であり、プライバシー等の倫理的な問題は一切生じない。また、2大学での実験的試用に際しては、性・年齢(年齢の詳細を明らかにしたくない場合には、99と入力)・学籍番号の情報をフェイスシート画面に提示し、同画面上部には、強制的調査ではないことを明記した。学籍番号は再テスト測定による複数回データの対応付けのためにのみ用い、解析時点では完全にその学籍番号のデータ部分を削除することを説明し、同意を得た。

C. 研究結果

1. MGRM 解析

因子分析(主因子法)の結果、第1因子の固有値は8.84(説明分散:36.8%)で、第2因子以降の固有値(分散%)は、1.76(7.3%)、1.20(5.0%)、1.04(4.3%)、0.96(4.0%)、0.87(3.6%)と推移していた。第1因子の固有値は第2因子のその5倍であり、Scree plotの適用でも明らかに1因子構造が支持された。

次いで、MGRMによるIRT解析を行なった。図2に全24項目の反応カテゴリ特性曲線を示す。『ポジティブ感情』4項目の各回答カテゴリ曲線は、#16を除いて、著しく平坦であり、 θ の変化と回答選択肢の対応が悪い(すなわち、鋭敏さに欠けている)ことが分かる。一方、それらのネガティブ表現改訂項目は、『うつ感情』と同様の高い峻別性を有している。

全体として、回答選択肢「めったにまたは全くない(1日未満)」と「いくらかまたは少しはある(1-2日)」の各カテゴリ曲線が交わる位置(b1)は潜在特性値 θ の中央付近であり、それより程度の重いカテゴリ間の交点b2およびb3が高値側にあることから、CES-DR尺度は比較的うつ症状の高いレベルを捉える尺度である

ことがうかがえる。このことは、尺度全体の情報関数曲線でも確認された(図3)。

2. 第1回目測定

表4左側にCATの終了条件とした θ の推定誤差0.30という収束条件までの所要回答項目数の分布を示す。平均は 10.4 ± 2.7 (6~15)項目で、約7割の回答者が11項目までで収束した。一方、37名(18%)はもう一つの終了条件の15項目までに収束しなかった。回答に要する時間は平均 48.5 ± 20.7 秒で、最短15秒~最長1分57秒でCAT測定を終えていた。全24項目に順に答えていくCBT測定では、回答所用時間は、平均2分1.9秒 ± 41.5 秒で、最短35秒~最長4分14秒かかっていた。

収束(提示)項目数ごとの回答者の平均 θ_{CAT} レベルを見ると、少ない項目で収束している回答者の方が、うつレベルが高くなっており、その傾向は11項目まで見られている。12項目から14項目での収束ケースは、それぞれ11項目までと比べて極端に少なくなっており、12項目で収束した9名の平均 θ_{CAT} 値は1.43と、この群だけが、再び高い推定値に戻っていた。最大提示数の制約15項目までかかった回答者の平均 θ_{CAT} 値は-0.76と、うつレベルのかなり低い群であった。

図4にCATの収束項目(横軸)ごとの平均 θ_{CAT} 値およびその回答者の平均 θ_{CBT} 値を示す。全体として θ_{CAT} 値と θ_{CBT} 値の対応は良好で、とりわけ11項目までは非常に高い一致が認められる。12項目以降の収束群ではばらつき(図中のバーはse)が大きくなっており、強制終了の15項目ではseは重なっていない。13項目以降の平均値を見ると、 θ_{CBT} 値に比べ θ_{CAT} 値の方がやや低め(すなわち、うつ症状が軽め)に推定される傾向が見られた。表5の上半分は1回目測定時の4測定値間の相関係数を示す。いずれの相関係数も0.90以上の高

値を示しており、 θ_{CAT} 値と θ_{CBT} 値の相関係数は 0.94 であった。

3. 第2回目測定

表4右側に2回目のCATの収束条件までの所要項目数の分布を示す。平均は 10.7 ± 2.9 (6~15) 項目で、約7割の回答者が12項目までに収束した。一方、39名(23%)は15項目までに収束しなかった。CAT測定に要する時間は、平均 33.4 ± 15.2 秒で、最短9秒~最長1分39秒であった。CBT測定の所用時間は、平均1分32.2秒 ± 32.9 秒で、最短46秒~最長4分14秒であった。

提示項目数ごとの平均 θ_{CAT} レベルは、1回目と同様に、少ない項目で収束している者の方が高かった。12~14項目収束者は、11項目収束者とほぼ同じ θ_{CAT} レベルであった。15項目で強制終了した回答者の平均 θ_{CAT} 値は-0.73と、第1回目と同様の結果であった。

図5にCATの収束項目ごとの平均 θ_{CAT} 値およびその回答者の平均 θ_{CBT} 値を示す。2回目も θ_{CAT} 値と θ_{CBT} 値の対応は良好で、特に11項目までは非常に高い一致が認められた。一方、12項目以降の収束群ではばらつきが大きくなっていた。13項目および14項目収束群では、 θ_{CBT} 値より θ_{CAT} 値の方がやや低めに推定される傾向が見られた。表5の下半分は2回目測定時の4測定値間の相関係数を示す。いずれの相関係数も0.92以上の高値を示し、 θ_{CAT} 値と θ_{CBT} 値の相関係数は0.95であった。

4. 再テスト信頼性

表5の対角部分は、4測定値それぞれの1回目と2回目の測定値間の相関係数(再テスト信頼性)を示す。いずれの相関係数も0.86以上の高値を示しており、十分な再テスト信頼性が認められた。

5. 回答時間測定 (CBT)

CBTの項目提示から回答までの応答時間を『身体症状』、『うつ感情』、『対人関係』、『ポジティブ感情』、『ネガティブ表現改訂』の5下位尺度ごとに算出した(図6)。『ポジティブ感情』への回答は、男女とも約7秒かかっており、他の項目よりも2秒以上の時間を要していた。一方、『うつ感情』は最も男女差が大きく、女子の方が有意に時間がかかっていた。『身体症状』でもその傾向が見られたが、他の下位尺度では男女差は認められなかった。

一方、回答時間が θ_{CBT} 値と有意な相関 ($p < 0.01$) を示した項目は、『うつ感情』の3項目#14「一人ぼっちで寂しい」、#17「急に泣き出したくなる」、#18「悲しいと感じる」、および『対人関係』の#19「皆が自分を嫌っていると感じる」で、『ネガティブ表現改訂』の#24「毎日が楽しくない」が有意傾向 ($p < 0.06$) であった。これらはいずれも正の相関であり、回答時間が長いほど θ_{CBT} 値が高くなっていた。

D. 考察

うつ症状を標的としたIRT-CATシステム(終了設定:推定誤差0.30未満)を大学生209名に試用した結果、11項目程度までで収束した。また、推定 θ_{CAT} 値 ≥ 1.5 の高うつ群のほとんどは7項目程度までで収束していた。所要時間も大幅に短縮していた。

1週間間隔での再テスト信頼性は十分に高い値を示した。CBTとの対応による並存的妥当性の検討でも、きわめて高い一致が認められた。これは、全24項目に回答しなくとも、ほとんどの回答者は半分程度の項目で、また特にうつレベルの高い群では1/3程度の項目に答えるだけで、うつレベルが推定できることを示している。

このように、本CATシステムの利点が確認される一方、回答者の1/4程度がその推定値精

度に問題があることが伺われた。これは本システムのもとになった CES-DR の測定特性(図3)から予め想定される事態であり、この尺度自体が低うつレベルの測定に対応していないことに起因する。この種のうつレベル評価の主目的は、うつ「高リスク状態」にある者の検出であるので、現行のシステムでもその目的は達成されるものと思われる。

ただ、現在までに報告されているメンタルヘルスに関する CAT システム (Fliege ら, 2005 ; Walter ら, 2007) は、既存の標準化尺度から 50 項目を超える豊富な項目プールを備えていることを見ると、今後の課題は自ずと明らかである。一方、妥当性については、最も重要な臨床的妥当性に関する検討が望まれる。

一方、一次元性が基本的条件である通常の IRT に基づく IRT-CAT では、測定評価領域が狭くなりすぎるのが懸念される。労働者のメンタルヘルス評価を考えると、今回のうつ症状のみならず、不安や身体愁訴など、個別に CAT 化した上で、複数の CAT (CBT も含む) を連結して実施する必要があるのかもしれない。本分担研究としては、過重労働等によるストレス負荷の測定法の中核部分である、うつ状態に焦点を当てた『精神的ストレス反応』モジュールの信頼性・妥当性が確認されたことで、次年度の総合システムの構築がより現実的になったと言える。

E. 結論

本分担研究課題の中心となるストレス反応測定評価のためのコンピュータを用いた適応型テスト (IRT-CAT) モジュールの検討のために、抑うつ尺度および追加項目 (計 24 項目) の CAT システムを構築し、大学生 209 名に試用した。うち 168 名には約 1 週間の間隔で再度回答を求めた。また、同一画面形式で全項目に回答を求める CBT も同時に実施した。

その結果、うつ測定 CAT の十分な再テスト信頼性と高い並存的妥当性が確認された。これにより、特にうつ症状に焦点を当てたストレス反応測定モジュールが整備され、次年度の総合システムの構築に向けた準備がほぼ整ったと考えられた。

文献

- Fliege, H., Becker, J., Walter, O. B., Bjorner, J. B., Klapp, B. F., & Rose, M. (2005). Development of a computer-adaptive test for depression (D-CAT). *Quality of Life Research*, 14, 2277-2291.
- Iwata, N. & Buka, S. (2002). Race/ethnicity and depressive symptoms: a cross-cultural/ethnic comparison among university students in East Asia, North and South America. *Social Science & Medicine*, 55, 2243-2252.
- Iwata, N. & Higuchi, H. R. (2000). Responses of Japanese and American university students to the STAI items that assess the presence or absence of anxiety. *Journal of Personality Assessment*, 74, 48-62.
- Iwata, N., Roberts, C. R., & Kawakami, N. (1995). Japan-U.S. comparison of responses to depression scale items among adult workers. *Psychiatry Research*, 58, 237-245.
- Iwata, N., Saito, K., & Roberts, R. E. (1994). Responses to a self-administered depression scale among younger adolescents in Japan. *Psychiatry Research*, 53, 275-287.
- Iwata, N., Umetsue, M., Egashira, K., Hiro, H., Mizoue, T., Mishima, N., & Nagata, S. (1998). Can positive affect items be used to assess depressive disorders in the Japanese population? *Psychological Medicine*, 28, 153-158.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied*