

PC キーボードの階層的操作性評価モデルの提案

富田 新

1 はじめに

キーボードをはじめとする電子デバイスのユーザビリティ（以下操作性と表記）の評価に用いられる指標は、評定尺度（質問紙）、筋電、ハイスピードカメラによる運動過程の解析、脳波など多岐にわたっている。多くのユーザビリティラボでも複数の指標を用いて多角的に操作性を評価するという方法が採用されているが（例えば、田村，1998）、指標間の相互関係を明確にした上で、系統的な評価が行われているかという点、必ずしもそのようにはなっていないようである。その理由としては、各指標の意味づけが必ずしも明確でないこと、指標間の関係が明らかになっていないこと、それぞれの機器操作で“心理学的・運動生理学的モデル”¹⁾が十分に構築されていないこと、などが挙げられよう。

特定の機器について有用な操作性評価システムを開発してゆくためには、機器を操作する際のユーザの“心理学的・運動生理学的メカニズム”を明らかにしつつ、種々の操作性評価の指標を適切に割り付け、指標間の関連性を総合的に解析してゆく必要があると思われる。

本論では、パーソナルコンピュータ（以下PCと略記）のキーボードの操作性を取り上げ、文章入力課題遂行時のPCキーボードの操作に関わる“階層的タスクモデル”を構築し、そのモデルに基づき、PCキーボードの新しい“階層的操作性評価モデル”²⁾を提案する。

本論で提案したモデルの着想は、平成18年度から平成19年度にかけて行われたアルプス電気（株）からの委託研究³⁾を推進する過程で得られたものである。この研究プロジェクトは、いわき明星大学・機械工学科の清水信行研究室（当時）を中心として行われた。PCキーボードの操作性を、感性評価（評定尺度を用いた質問紙）、生理指標（筋電、脳波）、指の運動解析、作業パフォーマンス等の複数の評価指標を用いて総合的に明らかにしようと試みるプロジェクトであった。本論で提案される操作性評価モデルがPCキーボードの操作性に特化しているのはそのためである。

本論において提案される“階層的操作性評価モデル”は、2009年に提出されたプロジェクトの最終報告書（清水，2009）において提案された“PCキーボードの操作性評価システム”に基づいている。システムの概要については日本応用心理学会第81回大会（2014：中京大学）においても報告された⁴⁾。

2 PC キーボードの操作性評価モデルの基本的な考え方

まずPCキーボードの新しい操作性評価モデルに必要とされる要件について検討を行った。以

下にそれらを列挙する。

- (1) PC キーボードの操作性の総合評価(「好ましい」、「操作性が良い」等)を予測するためのモデルを構想する。
- (2) そのためには、PC キーボードを用いて行われるタスク(本論では“文章入力課題”を想定した)を分解し、キーボードの操作性の総合評価に関わると想定される種々の下位タスクを特定してゆく必要がある。また、それらのタスク間の関連性をある程度体系的に記述してゆく必要がある(キー操作の“階層的タスクモデル”の構築)。
- (3) それらのタスクに対して種々の評価指標を割り付ける。そのためには、
 - ①個々のタスクの評価指標、及び、その測定法・分析法・評価法をある程度確立しておくことが必要。
 - ②評価指標間の関連性を把握し、モデルに組み込んでおくことが必要。
- (4) 個々の下位タスクの評価指標がどのようなパターンを示せば、総合的なキーの操作性が「良い」(もしくは「悪い」と評価されるのか、をモデル化する(各指標の“重み付け”等を考慮する)。
- (5) キー操作への慣れや学習の程度、評価に現れる個人差などを、ある程度予測できるモデルであればなお望ましい。キーの操作性評価の個人差は、キー操作への慣れや学習の過程で生じる各下位タスクの評価の個人差や、各下位タスクへの“重み付け”の相違(もしくは変化)として捉えられる。
- (6) 各タスクをさらに下位のタスクへと分解しモデルを詳細化できる、新しい評価指標を割り付けてモデルの構造を新たに組み替えられる、など、柔軟性・発展性をもたせたモデルとする。例えば、使用される機器の種類や課題遂行の手順の変化等にも柔軟に対応できるモデルとする。また、新しい評価指標等が開発された場合は、新たな指標を位置づけてゆくための準拠枠(フレームワーク)として参照できるようにする。

3 ユーザのキー操作モデル(階層的タスクモデル)をベースに操作性評価モデルを構築する理由

新しい操作性評価モデルを考案する際に、ユーザのPCキーボードのキー操作の“心理学的・運動生理学的モデル”(以下“階層的タスクモデル”と表記)をベースとすべきであると考えた理由は以下の通りである。

- (1) 文章入力時のキーボードの操作は様々な下位タスクの複合体であり、マクロなレベルからミクロなレベルまで、幾多の階層性を有する。例えば、PCキーボードを用いて文章入力課題を遂行する場合を想定した場合、最もマクロなレベルでのタスクは、“文章入力”であるが、それを遂行するためには、“文字キーの特定”、“指の上げ”、“キーの押下”、“クリックの感知”、“再度の指上げ”など、種々のミクロなタスクが滞りなく行われる必要がある。幾つかの評価指標(例えば、特定の筋肉から得られる筋電や1個のキースイッチに対する押し心地評価など)は、主にミクロなタスクを評価するための指標である。しかし、これらのミ

クロな指標をばらばらに見ているだけでは、総合的な操作性評価がどのような過程を経て生じているのか、また、機器のどのような特徴が総合的評価を下げているのか等を明確に特定できない可能性がある。各評価指標間の関係を踏まえた上で、キーの操作性を総合的に捉える視点が必要であると思われる。機器に対するユーザの総合評価（好み・商品選択の判断など）は、それら個々の下位タスク評価を総合した結果として出力されている可能性が高いと思われる。

- (2) (1) の考え方にに基づき総合的な操作性評価モデルを実現するためには、複数の操作性指標を総合し、指標間相互の関係性を記述するための土台（フレームワーク）となるモデルが必要になると考えられる。その際、何を土台にすべきかが重要な問題となるが、本論では“ユーザの機器操作過程そのものを土台とすべき”と考えた。つまり、ユーザがPC キーボードを使用して“文章入力課題”を行う際の“心理学的・運動生理学的モデル”こそが、複数の操作性指標を総合的に扱えるようにするためのフレームワークとなり得る、と考えたのである。本プロジェクトで使用された操作性評価の指標（評定尺度を用いた質問紙、筋電、脳波、ハイスピードカメラによる運動過程の解析等）は全て、この“心理学的・運動生理学的モデル”のどこかに必ず位置づけられることになる。このとき注意すべきは、各指標の捉えている機器の操作性が、PC キーボードの“好み”といった心理的・抽象的なレベルから、“指上げ・指下ろしのしやすさ”といった運動生理学的・具体的な反応まで多岐にわたっている、ということである。従来は、タスクのレベルの違いや階層性についてはあまり考慮せずに、各指標による測定結果を並列的に記述することによって、機器の操作性評価が行われることが多かった。本論のモデルでは、階層の異なるタスク遂行を想定し、階層間のタスクの関係性を明確に記述しようと試みる。例えば、①ユーザの課題遂行意図という“構想のレベル”、②①の意図を実現するためにユーザの内部に生成される操作手順・操作計画などの“計画のレベル”、③実際に感覚器や効果器によって課題（機器操作）が遂行される“実行のレベル”という3つのレベル（階層）⁵⁾を想定し、それぞれの階層に配置されたタスクに分解して、各階層のタスクに対応する操作性指標を割りつけてゆくことを考えるのである。このような考え方は脳科学における運動生理学の知見とも矛盾しない（例えば、入来，2011）⁶⁾。また、ユーザの課題遂行の“心理学的・運動生理学的モデル”に基づき、操作性評価モデルを構築することは、“ユーザの機器操作の心理学的・運動生理学的メカニズムの検証”という基礎心理学的な観点と、“特定の機器の操作性評価”という応用心理学的な観点とを結びつけることにもつながる。

4 キー操作の階層的タスクモデルの概略

以下に、ユーザのキー操作の階層的タスクモデルの概略を述べたい。

4.1 キー操作の階層構造とタスク分解

キー操作のタスクを4階層（“文章入力タスク”、“中間タスク”群、“キー入力操作タスク”群、

“指の上げ・下ろしタスク”群)から捉えた(図1)。

まず、最高位(第1階層)のタスク(“文章入力タスク”)を想定する。この最高位のタスク遂行において喚起される“構想”(入来(2011))または“目標”(甘利(2007))を実現するために、機器の使用が行われ、機器の操作性の総合評価(評価A:例えば、機器の好悪の評価等)が行われると考える。

より上位(第1階層)のタスクを実現するために、幾つかの“中間タスク”群(第2階層のタスク群)が計画され、系列的に配置され、遂行される(“文章の読み取り”、“キー入力操作”、“モニタによる入力確認”)。これらのタスクは何度も繰り返される。すなわち、ループを形成する。エラー修正のためのループもその中に組み込まれる。“中間タスク”(第2階層のタスク)のそれぞれについても、目標が形成され、その目標に対して結果が絶えず評価される(評価B1、評価B2、評価B3)。これらの“中間タスク”の遂行が1つでもスムーズになされなければ(例えば、“モニタのコントラストが不明瞭”、“キーが探しにくい”、など。このときB2、B3の評価は“-”(マイナス)となる)、それは、1階層上のレベルのタスク(第1階層のタスク)の評価(この場合、“文章入力タスク”の評価A)を下げる方向に働く。

第3階層はよりミクロなレベルのタスクから構成される。図1では、“キー入力操作タスク”の部分のみをクローズアップして図示しているが、他の“中間タスク”(第2階層のタスク:例えば、“モニタによる入力確認タスク”)についても、よりミクロな下位タスク群(第3階層、第4階層・・・のタスク群)を想定することが可能である。図1では、“キー入力操作タスク”の下位タスクとして、“キーの同定タスク”、“キーの押下タスク”、“クリックや反発性の感知タスク”、“再度の指上げタスク”を想定している。これらのタスクも何度も繰り返され、ループを形成する。また、エラー修正のためのループも組み込まれる。これらのタスクに対しては、目標(“内部モデル”または“運動プログラム”:甘利(2007))が形成され、目標に対する結果が絶えず評価されるものとする(評価C1、評価C2、評価C3、評価C4)。

これらの下位タスクの遂行が“-”(マイナス)評価であるときには、より上位のタスク(“キーの入力操作タスク”)の評価(B2)は低下するものと仮定する。その結果、さらに上位のタスク(“文章入力タスク”)の操作性評価(総合評価)(A)も必然的に低下することになる。つまり、ある階層にボトルネックとなるタスク(“-”評価を受けるタスク)が存在する場合には、より上位のタスクの評価が全て“-”方向に引きずられることになると想定する。

第4階層は今回提示されたモデルにおける最もミクロなタスク群である。図1では第3階層の“キーの押下タスク”の下位タスク(第4階層のタスク)として“指上げ”タスクと“指下ろし(押し込み)”タスクを想定している。これらのタスクの評価が“-”になるとき、より上位の階層のタスクの評価が全て“-”方向に引きずられることになるとは上に述べた通りである。

以上のようなタスク分解を行う際には、心理学的観点や感覚・運動生理学的観点から見て、各タスクの独立性(モジュール性)がある程度保証されていることが必要である。図1のタスク分解の妥当性については、現時点では必ずしも確証が得られているわけではない。そのため、タスクの分解の仕方が今後変更になることも十分にあり得る。

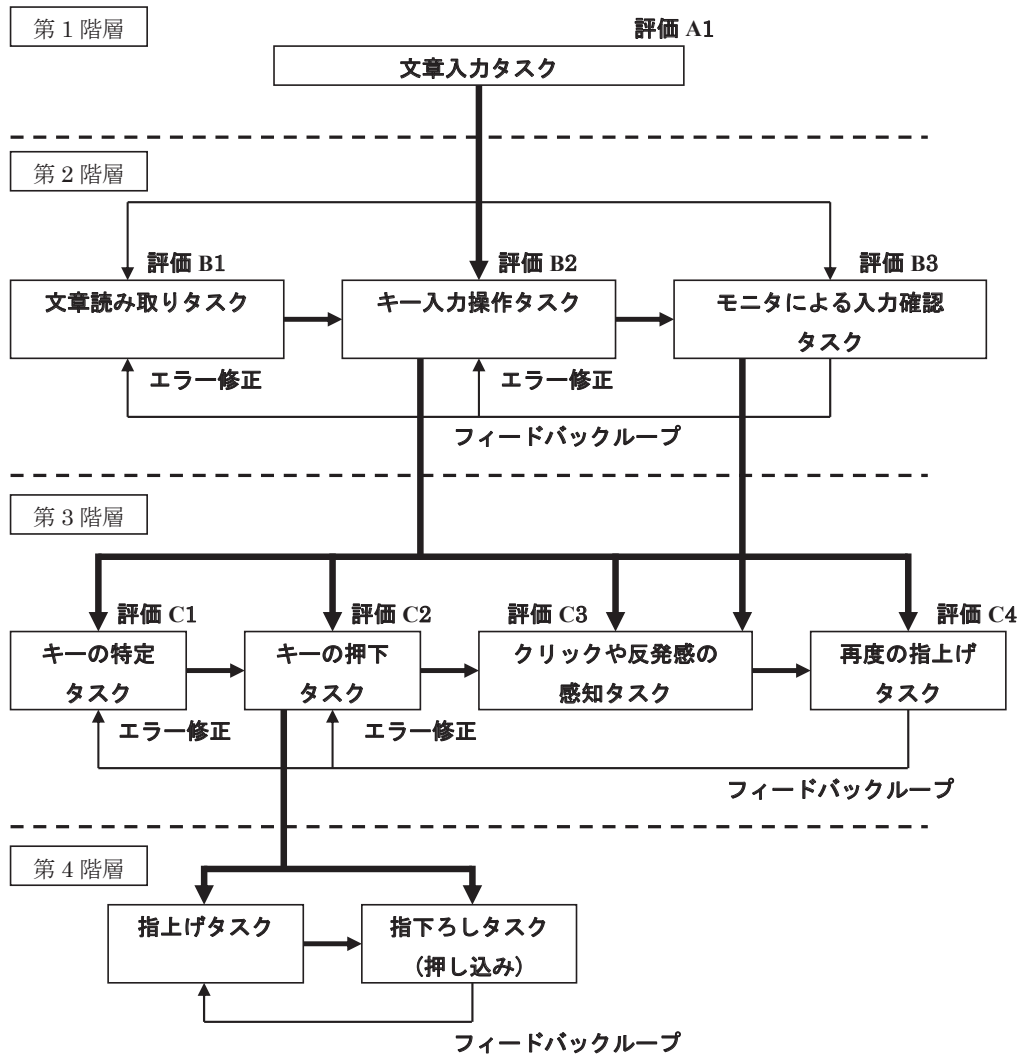


図1 キー操作の階層的タスクモデル

4.2 目標・予測に基づくタスク遂行の評価

個々のタスク遂行の評価においては、以下の基本仮説を設定する。

キー操作の各タスク遂行時に生じているユーザの目標（必ずしも意識化されているとは限らない）が、ユーザの予測通りに（なるべく予測に近い形で）実現される場合（すなわち、ユーザの予測したタイミングで、予測した効果を生じさせ、予測した通りのフィードバックが行われる場合）、また、その際、ユーザにかかる物理的・心理的負荷が少ないとき、そのキーは「操作性が良い」と評価される確率が高くなる。一方、それらのタスクのいずれか1つでも、ユーザの予測からずれる場合には、不快感や違和感が生じ、「操作性が悪い」と評価される確率が高くなる。

4.3 キー操作の“学習”・“慣れ”と評価の“個人差”について

キー操作時に最初に形成される目標は、機器の特性に合わせて異なってくる。指の動きなどの筋骨格系（効果器）の動きは、基本的に、中枢の“意図”や“内部モデル”に基づき発動される。最初に生成された目標に対して効果器がもたらした結果が感知され、中枢にフィードバックされる。フィードバック過程で最初の目標と感知された効果の差分（“報酬予測誤差”（例えば、中原(2007)）と呼ばれているような予測と結果の差分。以下“予測誤差”と表記。）が計算され、評価され、記憶される。繰り返し操作によってこの差分の記憶が安定したとき、その後の操作の評価基準として作用するようになる。つまり、その後の操作では、記憶された差分を基準として、目標の設定、予測、評価等がなされるようになる。

例えば、キーボード操作では、通常“キーの押下”の目標が形成される。その目標に応じて“キーの押下”の動作が行われるが、このとき、キースイッチのストロークがなく、キースイッチの沈み込みが確認されない場合、最初に生じた目標と感知された結果の差分（“予測誤差”）が大きくなる（この場合、差分は負（-）の値をとると仮定）。この“予測誤差”が、記憶された基準に照らして評価され、“違和感”や“不快感”の喚起（あるいは“操作性の低さ（悪さ）”の感知）につながる。“予測誤差”の方向性（+、-）によって操作性の評価がなされる。+の“予測誤差”はその操作法をより強める（強化する）方向に、-の“予測誤差”はその操作法を修正する方向に作用する。なお、強化学習モデルでは、“報酬予測誤差”をなるべく0（ゼロ）にする方向に学習が進むと想定されている（中原, 2007）。

しかし、この違和感（-）を伴うフィードバックが何度も繰り返されると、目標そのものが、フィードバックにもとづいた修正を受ける。この繰り返し（フィードバックループ）により、中枢での目標（内部モデル）も微妙に修正され、それに基づき効果器の動きも修正される。最終的に、その機器の特性に合わせた効果器の動きが動員されるようになる（“慣れ”や“学習”の過程）。

例えば、“ストロークのないキーの入力”では、キーを“押し込もうとする”（または“叩く”）ような指の動きではなく、例えば、キーを“触る”ような動きへと、操作に伴った運動が修正される⁷⁾（“学習”や“慣れ”によって“操作の目標”と“操作の仕方”が変化する）。

また、キー入力時の効果器の動き（指の動かし方）には、元来個人差があるため、個々人で異なった目標（個々人の指の動きの特性に合わせた“内部モデル”や“運動プログラム”）が形成され、異なった効果器の動きがもたらされ、異なった結果がフィードバックされる。最終的に、個々人で異なったキーの操作性評価が行われるようになる。この異なったフィードバックループが繰り返された結果、キー押しの仕方やキーの操作性評価の“個人差”が生じると考えられる。つまり、機器の特性と個人の特性（体型など）の相互作用により、キー操作時に形成される目標（内部モデルや運動プログラム）や操作性評価の基準が徐々に変化してゆき、最終的に個々人の操作法や評価基準が安定してゆくと考えられる（キー操作の仕方や操作性評価の“個人差”が形成される）。

4.4 “クリック”の重要性について

キー操作における“クリック”（“ONに関するフィードバック”⁸⁾）は、キー操作において特異かつ重要な意味をもつ。すなわち、“必須である”と想定する。なぜなら、通常キー押しはこの“ON”

を目標としてなされるからであり、中枢で形成される目標に対して“ON”の感知は最も重要な信号となると予想されるからである。したがって、“ON”に関するフィードバックが得られない状態というのは、ユーザに極めて強い違和感や不快感を生じさせ、キーの操作性を著しく損なう可能性がある。また、“クリック”のタイミングも極めて重要になる。ユーザの予測にできるだけ沿ったタイミングが望ましい。

ただし、“クリック”（“ONのフィードバック”）は必ずしも“タクトイルな”（“触知的な”）感覚モダリティを通して得られる必要はなく、他の感覚モダリティの刺激（例えば、“聴覚刺激”や“視覚刺激”）でもかまわない可能性がある。そのため、図1においては、“キー入力操作タスク”（第2階層）の遂行に伴って形成される第3階層の下位タスクのフィードバックサイクルの外側（例えば、より上位の階層にある“モニタによる入力確認タスク”など）から“ONのフィードバック”がかかることを想定した（図1参照）。

また、ユーザが、異なる複数の“クリック”（例えば、“キースイッチから得られるタクトイルなクリック”と“モニタ上に提示される視覚的クリック”）を感知し、利用することも考えられる。このとき、感知された複数の“クリック印象”（または“クリック感”）が、ユーザのタスク遂行の過程で、どのように統合され、活用されるかが問題となるが、詳細については今後さらなる検証が必要である。

キーボード操作では、その指の運動の複雑さ（それは、“キーの同定”とも密接に関わっている）から、タクトイルな“クリック”（キースイッチに機械的にセットされたパラメータとしては“ドロップ荷重”）が、聴覚的または視覚的なクリックとは独立に、重要な役割を果たしている可能性がある⁹⁾。もしそうだとすれば、安易にそれを無くすことはできないということになる。また、現在のキーボードに実装されている“クリック”（“ドロップ荷重”）は、腱鞘炎の防止という医学的な見地からも必須であることが知られている。

5 キー操作の階層的タスクモデルに基づいた新しい評価モデル（階層的操作性評価モデル）の提案

キーボードの新しい操作性評価モデルは、図1の“階層的タスクモデル”の各タスクに、適切な評価指標を割り付けてゆくことによって構築される。その際、各評価指標が、キー押しの各階層・各タスクのどういった側面を評価しているのかを十分に精査しておく必要がある。また、それらの指標の関係性についても、可能な限り明らかにしておく必要がある。

5.1 本プロジェクトにおいて採用された操作性評価指標の特徴

今回のプロジェクト研究において実際に使用された操作性評価指標の特徴と問題点を以下に挙げる。

(1) 感性評価（評定尺度を用いた質問紙）

意識的な操作性の評価指標である。「評価に主観が混入する」、「再現性に疑問が残る」等の問

題はあるが、感性評価(評定尺度を用いた質問紙)は多くの操作性評価において用いられている比較的ポピュラーな指標の一つである。各尺度は“機器の操作性に関する質問項目”と“評価スケール”の組み合わせによって構成される。これらの感性評価(評定尺度を用いた質問紙)で測定されたデータを分析することにより、尺度をさらに有用なものへと洗練させてゆくことができる。本プロジェクトでは、キースイッチにセットされた差動力曲線の各パラメータと心理評定との対応関係の分析(“重回帰分析”による線形モデルの抽出)や、エキスパートとノービスの評価傾向の違いについての分析(“判別分析”による線形モデルの抽出)が行われた。これらの分析結果に基づき、例えば、「エキスパートが重視する操作性評価指標としてはどの尺度(項目)を重視すれば良いか」等について、新たな指針を得ることができた(富田・大原(2015))。このような指針は、感性評価(評定尺度を用いた質問紙)の新しい使用法のバリエーションの1つであり、新たな評定尺度と成り得るものである。

感性評価(評定尺度を用いた質問紙)を用いる際に注意しなければならないのは、個々の指標が“階層的タスクモデル”(図1)の異なる階層の、異なるタスクに対応している可能性があるという点である。例えば、今回のプロジェクトで用いられたキースイッチの押し心地に関する評定項目の中では、「反発性がある」や「キーが柔らかい」は第3階層の“キーの押下タスク”等の評価(C2、C3)に対応していると考えられる。これに対し、「楽(押しやすさ)」は、第2階層の“キーの入力操作タスク”の評価(B2)に対応する評定項目であると考えられる(富田・大原(2015))。つまり、同じ1つの質問紙の中に組み込まれた評定尺度であっても、それぞれが異なる階層の、異なるタスクの評価指標として位置づけられることになる。

こういった感性評価(評定尺度を用いた質問紙)をより洗練させてゆくためには、①多変量解析等による統計解析を行い、評定尺度間の関係性を明らかにしてゆく、②実験的手法を導入し、キーの物理パラメータの変化と心理評定との関係をより実証的に把握してゆく等の作業が必要になると思われる。

(2) 課題成績

機器を使用し、特定の作業をさせたときの、“作業パフォーマンス”(“エラー率”や“作業(反応)時間”)である。客観的な行動指標であり、操作性評価の指標としての信頼性は高い。ただし“心地よさ”のダイレクトな指標では必ずしもない。それ故、感性評価(評定尺度を用いた質問紙)など、他の指標との関連性を踏まえた上で、使用を検討した方が良い場合もあると思われる。また、機器に対する“慣れ”や“学習”等の影響を無視できないという点にも注意を要する。機器への“慣れ”や“学習”が生ずることで、課題成績は通常上昇する。そのため、課題成績を操作性評価の指標として用いる際には、機器操作への“慣れ”や“学習”のどの段階で測定されたデータなのかについても把握し、考察してゆく必要があるだろう。

(3) 指の運動の解析

客観的な行動指標であるが、キー押しにおいては、解析の視点や評価軸がまだ十分に開発できていないという問題が残る。本プロジェクトではハイスピードカメラを用いた指の運動過程の撮

影と解析が行われた。しかし、打鍵法の個人差が大きく、有用な評価軸を見つけるには至らなかった。本プロジェクトの結果からは、キーの打鍵における“指上げの高さ”や“キーの押し込みの深さ”や“押し込みの持続時間”等が、有力な指標として利用できる可能性が示唆されたが、今後さらなる検証が必要であろう。

(4) 筋電

生理指標であり、正確に測定が行われれば、評価指標としての客観性・信頼性は高い。また、過去の知見から、評価指標としての位置づけも比較的明確であるといえる。例えば、筋電の量については、少ない方が“楽”であり、“疲労が少ない”とされる。また、操作中の状態の変化をリアルタイムに近い形で追えるというメリットもある。ただし、キー操作の“心地よさ”を反映するかどうかについては必ずしも定かではない。また、筋電の処理法・分析法にもまだ開発の余地がある。例えば、筋電のこういったパラメータ（電圧、持続時間、積分値等）を利用するかについては、明解な指針が得られているわけではない。

本プロジェクトで行われた一連の実験では、筋電に関して、2つの成果が得られている。

1つはストロークのないキーボードへの順応もしくは学習の過程を、筋電の変化によってある程度実証できたことである¹⁰⁾。分析結果から、ストロークのないキーボードに慣れることにより、被験者の“総指伸筋”の筋電が大きくなるとともに、“浅指屈筋”の筋電が逆に小さくなることがわかった。これは、ストロークがないキースイッチの特性に対応して、被験者がキーの押し方（叩き方）を変えている可能性を示している。ストロークがないことをカバーするために、「指を高く挙げ、振り下ろしの際の押圧を弱めている」と推察されるのである。ただし、このような操作法がストロークのないキーボードにおける一般的かつ最終的な操作法かどうかは、今回の実験結果のみをもって結論づけることはできない。ストロークのないキーボードへの順応期間（もしくは学習期間）が2週間と短いことや、全ての被験者で同様な傾向が見られたわけではないという点を考えると、これらの筋電の変化は一部の被験者の過渡的な操作法の変化を示しているに過ぎない、という解釈も捨てきれない。また、これらの筋電のパターンと、掌や腕の保持（姿勢維持）がどのように連動しているかについても、よくわかっていない。「指を高く上げ、押圧を弱めるような」動作が、掌や腕の姿勢保持に負荷的に作用している可能性は当然考えられる。それ故、ストロークのないキーボードが、ストローク有りのキーボードと同等な機器となり得るかどうかについて、本プロジェクトの実験結果のみから安易に結論付けることはできないと言える。

本研究で得られたいま1つの成果は、筋電データに基づき測定された“キー押しの収束時間”が他の様々な評価指標と有意な相関をもっていたということである¹¹⁾。その背景にどのようなメカニズムが働いているのか、いまのところ明らかではないが、“キー押しの収束時間”が有力な操作性評価の指標として利用できる可能性はあると思われる。今後より詳細な検証が必要であろう。

(5) 脳波

脳波は脳の活動状態をある程度反映すると考えられている生理指標の1つである。本研究で採

用した背景脳波 (β/α 値) は、脳波に含まれるパワースペクトルの α 波成分と β 波成分の比率であり、ユーザビリティ評価の指標として比較的良好に用いられている指標である (例えば、上野ら (2002))。

ただし、背景脳波 (β/α 値) には、操作性の評価指標として以下のような問題があることも事実である。

- ①電子機器操作時には、背景脳波 (β/α 値) にアーチファクトが混入することが多いため、リアルタイムな機器操作の評価指標として使いにくい。多くの研究 (例えば、上野ら (2002)) では、機器使用後、一定時間、安静・閉眼状態で測定した背景脳波から β/α 値を算出している。本研究では、機器操作後だけでなく、機器操作前の β/α 値も併せて算出し、 β/α 値の操作前後での変化パターンを見るという手法を採用した。しかし、機器操作時のリアルタイムな背景脳波ではなく、操作後の背景脳波から算出されることもあって、 β/α 値が、機器操作のどのような側面の評価を反映しているかについては疑問が残る面もある。「タスクによる脳波の変化は 60 秒から 100 秒の間持続するので、タスク後に脳波を測ることによって、タスクによる脳波の変化を見ることができる」との主張もあるが (Oohashi ら (2000) : 上野ら (2002) より引用)、本研究の分析結果からは、課題終了後の 60 秒間に β/α の値が相応に変動することが確認されている。よって、背景脳波 (β/α 値) の機器操作の評価指標としての信頼性は必ずしも確認されていないと言える。
- ②脳波に影響を与える要因は機器の操作性以外にも種々想定し得る。例えば、被験者の気分や思考などがそうであるが、これらについて厳密に統制することは難しかった。この点も、脳波が“多義的”で、“あいまい”であるとの批判を受けやすい原因の1つであろう。これは、脳波が脳内処理のどのような側面を反映した現象か、まだよくわかっていないところにも起因していると思われる。
- ③脳波の波形や、 β/α 値は、個人間変動 (個人差) 及び個人内変動が大きい。この変動が大きいことも、脳波が操作性の評価指標として使いにくい原因の1つである。複数の被験者の算出される β/α 値の平均値は、種々の変動要因が混入するために、あまり安定しない。そこで、 β/α 値を用いる場合、各個人の β/α 値の平均値をベースラインとして算出し、この平均値で各被験者・各条件の β/α 値の値を割ることによって、正規化を行うという手法がよく用いられている。例えば、上野ら (2002) の研究でも同様な処理が行われている。本プロジェクトでは、この正規化は行わなかった。この点も、本プロジェクトで、背景脳波 (β/α 値) による分析結果が曖昧となった原因の1つであると思われる。また、脳波については、今回用いた背景脳波 (β/α 値) のみではなく、1/fゆらぎなど、様々な評価指標が開発されている。今後はそういった指標も併せて、操作性の評価指標として脳波が利用可能かどうかについて、検討を加えてゆく必要があるだろう。

5.2 キーボードの階層的操作性モデルの提案

5.2.1 各指標の階層的タスクモデルへの割り付け

図1の“階層的タスクモデル”に、本プロジェクトで用いられた上述の操作性評価指標を割り付けてゆく。

例えば、第4階層の“指下ろしタスク”の評価には、“ハイスピードカメラによる指の運動過程の解析”や“筋電測定”のデータを割り付ける、また、5.1(1)でも述べたように、たとえ“感性評価（評定尺度を用いた質問紙）”による評価であったとしても、単独のキースイッチの押し心地評価に関する質問（例えば、「底付き感がある」）は、よりミクロなタスク（例えば、第3階層の“キーの押下タスク”）の評価指標とみなすべきであると思われる。

第1階層の“文章入力タスク”は、長時間キーボードを使用した際の“課題成績”（“エラー率”、“作業時間”）や“疲労”、“操作性の総合評価の質問項目”（例えば“キーボードへの好悪の評定”）等によって評価される。また“筋電の積分値”も、“疲労”との相関の高さから考えれば、“文章入力タスク”の評価指標として使用することが可能であると思われる。

“脳波”が、どの階層のどのタスクを評価する際の指標として最も妥当であるかについては、今のところ明確な指針は得られていない。しかし、“キー操作前後の背景脳波（ β/α 値）”は、第1階層の“文章入力タスク”の総合評価と関連している可能性は高いと思われる。“脳波”が操作性の総合的評価指標として期待される理由は、タスク遂行時の“快-不快”の程度をある程度反映すると考えられている点である。

以上のようにして、図1の“階層的タスクモデル”に、個々の評価指標を割り付けたものが図2である¹²⁾。本プロジェクトの研究結果に基づき、現時点で考えられる“階層的操作性評価モデル”を想定したものが図2ということになる。ただし、図2は、あくまでも本研究の結果に基づいた暫定的な“階層的操作性評価モデル”の素案（コンセプト：概念図）に過ぎず、今後は新たな評価指標を加えるなどして、モデルの具体化・詳細化を図ってゆく必要があるだろう。

図2のような“階層的操作性評価モデル”を、より具体的・実用的な操作性評価モデルへと近づけてゆくためには、

- ①各階層のタスクに割りつけられた操作性指標の信頼性の高いデータを確保する。
- ②各タスクの評価指標として複数の指標が存在する場合には、実証的データに基づき、
 - (i) 最も妥当性・信頼性の高い指標を選択する
 - (ii) 複数の指標を統合した新しい操作性評価指標を考案する
 - (iii) 各指標を投入したときのモデルの適合性を検証し、適合性が最も高くなる指標を選択するのいずれかの方法に基づき、評価指標の選別を行う。
- ③これらのパラメータを投入し、抽出された数理モデルの交差妥当性や予測妥当性を検証し、最終的に、最も実用的で予測性が高いと考えられる数理モデルを採択する¹³⁾。

等の手続きが必要になろう。

本論で提案された“階層的操作性評価モデル”を、より具体的で実用的な操作性評価モデルと

して発展させてゆくためには、データに基づいた最適な数理モデルの抽出と、抽出された数理モデルの検証の手続きが不可欠になると思われる。この点については、特に強調しておきたい。

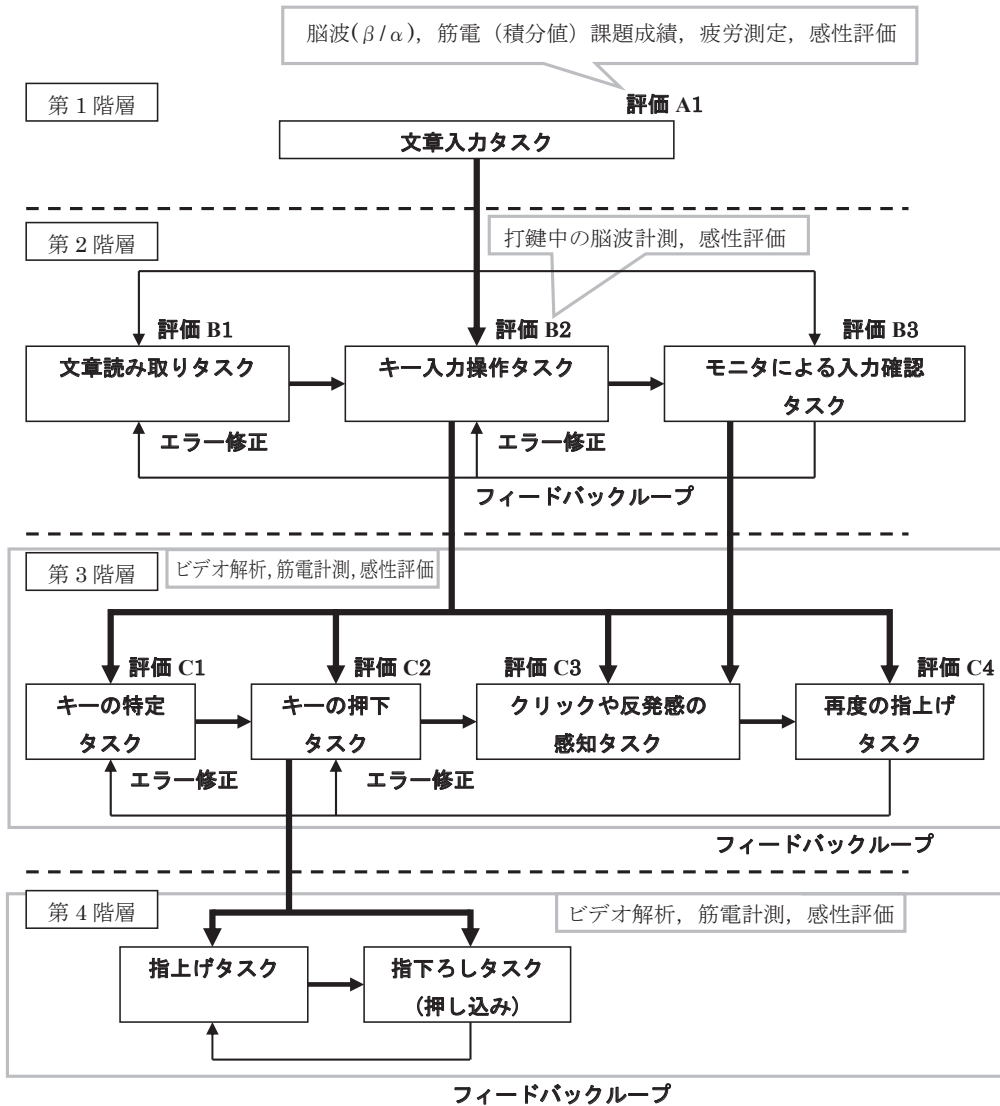


図2 本研究で使用された操作性評価指標に基づいた階層的操作性評価モデル

5.3 上記評価システムの問題点と今後の課題

(1) タスク分解の妥当性

図2の評価システムは、“文章入力タスク”を独立性の高い下位タスクに分解し、個々の下位タスクに評価指標を割付け、タスクベースで操作性評価を行うこと意図している。しかし、下位タスクを図1（または図2）のように分解したことの妥当性については、現時点では、必ずしも実証的根拠があるわけではない。例えば、第4階層のタスク（例えば、“指下ろしタスク”）を、動作解析や運動生理学の知見に基づき、より詳細な下位タスクに分解することも十分に可能であると思われる。いずれにせよ、タスク分解においては、各タスクのモジュールとしての独立性が心理学的・運動生理学的観点から見て保証されていることが必要である。

(2) 動的なタスクと静的なタスク（姿勢維持タスク）との関連性

キー操作において、静的なタスク（主として姿勢維持タスク）と動的なタスク（例えば、連動した指の運動）がどのように関連するのかが、図1のモデルには書き込まれていない。静的なタスクを評価するためには、まずそのタスクを特定し、新たに評価指標を割り付けていく必要がある（例えば、姿勢を維持している腕や肩の筋肉の疲労測定など）。キー操作に伴う静的なタスクの特定は、今後図1のモデルや図2の評価システムをさらに精緻化していく上で重要な課題であると思われる。

(3) 問題（ボトルネック）となるタスクの特定

図2の“階層的操作性評価モデル”は、“文章入力”という課題遂行において生じる個々のタスク評価を手がかりに、どのタスクに、どのような“予測誤差”（目標と感知された結果の差：“違和感”や“不快感”、あるいはタスクの流れを妨げるような“障害”）が生じているかを探索するためのツールとしても提案されている。一連のタスク遂行においてボトルネックとなっているタスクを特定することができれば、その解消に向けて、様々な対応をとることが可能になる。ただし、あるレベルのタスクで“予測誤差”が生じていることを特定できたとしても、その“予測誤差”の真の原因（ボトルネックとなっているタスク）の特定には、すぐにはつながらないことも十分にあり得る。特に、上位の階層のタスク評価は抽象度が高いため、そこから真の原因を特定してゆくことが難しい場合もあると思われる。真の原因（ボトルネック）となっているタスクの特定を容易にするためには、評価指標間の関係把握をしっかりと行い、モデルの精緻化（各指標を総合する数理モデルの精緻化と適合化）を図っておく必要があると思われる。

タスク遂行の流れを妨げていると思われる要因（ボトルネックとなっているタスク）については、評価を下げる方向に働く可能性があるため、機器の改良・開発においてできるだけ排除する。例えば、キー押し開始時に大きな力が必要とされる場合には、キーが重く感じられ、“キーの押下”タスクの評価を下げる方向に作用すると考えられる。その結果、キーボードの“総合的操作性評価指標”も低下することが予想される。“階層的操作性評価モデル”を用いた数値的解析の結果から、“キーの押下”がキーボードの総合的操作性評価指標を下げていること（“因果連鎖”）を

特定することができれば、ボトルネックとなる要因(この場合、キーの作動力の初圧の高さ)を排除することで、キーボードの総合的操作性評価を高めることができると予想される。

(4) タスク系列の制御の問題

幾つかのタスクは、一連のカスケードとなって自動的・系列的に進行していくものと思われる。このフィードフォワード制御において、ユニットとして、どのようなタスクがプログラムされているのか、また、目標と感知された結果の差分(“予測誤差”)が、どのタスクに対するフィードバック効果としてもたらされているか、さらには、そのフィードバックに基づき、どのように運動プログラムの修正がなされているのか、等については、図1のモデルでは、必ずしも明確に述べられているわけではない。また、運動制御の分野で問題視されている、関節などの自由度を減少させるための脳内の制御過程(例えば、“Bernstein問題”(Bernstein, 1967))についても未回答のままである。この問題に関しては、中枢神経系が同一のタスク遂行に参与する複数の筋を、まとまりとしてコントロールすることで自由度を減らしているのではないかという“シナジー”の考え方が提案されている(d'Avella et al, 2002, 2005, 2006))。これらの問題に回答を与えるためにも、キー操作に関する精緻な“心理学的・運動生理学的モデル”を構築する必要があるものと思われる。

(5) 機能性評価と質感評価(キーの操作性評価の“一般性”と“個人差”)の相違

従来機器の“機能性評価”と“質感評価”は別物として捉えられてきた。これらの評価が図2のモデルの中でいかに記述され、把握されるのか、すなわち、機器操作における“一般性”と“個人差”の問題が、図2のモデルの中でどのように切り分けて扱われるのか(そもそも切り分けて扱えるのかどうか)ということについても、詳細は不明のままである。将来的には、①機器の一般的な操作性(“機能性評価”)を規定する各パラメータの値と、②個々人の操作性の“好み”等を規定する各パラメータの値とを切り分けて、機器として備えるべき標準的な操作性(“機能性評価”)と、各ユーザの好みを反映する操作性(“質感評価”)をできるだけ正確に予測できるモデルを構築することが目標となろう。ユーザの機器の選択は、あるレベルまでは機器の“機能性評価”に基づいて行われている可能性があるが、あるレベルからは、ユーザごとの異なる操作法や評価基準(“質感評価”)の影響を受けて、個々のユーザごとに異なってくるのが予想される。ただし、1つの操作性評価モデルから、これらを2つの評価を切り分けて把握することが可能なかどうかについては現時点では判然としないところがある。少なくとも、図2のモデルのさらなる精緻化、信頼におけるパラメータの特定、複数のユーザのデータから抽出された信頼の一般性のある数理モデルの抽出、各個人のデータから抽出される数理モデルと一般的な数理モデルの比較・検証等が必要になるだろう。

(6) 多様な指標の活用(fMRI、PETなど)

昨今は、fMRIやPET等、脳内の機能的活動を測定しイメージングする技術が急速に発展しつつある。この種の、より高次元指標を用いた研究知見に基づき、図1や図2のモデルをさらに精緻化し、発展させてゆくことも今後の課題の1つである。

(7) キー操作に関わる基礎研究と応用研究の融合

キー操作に関わる研究を発展させてゆくためには、より普遍性のあるキー操作の機能モデル(脳科学や運動生理学、神経計算科学等の知見を含めた運動制御や運動系列学習等のモデル：本論で言うところの“心理学的・運動生理学的モデル”)を明らかにしてゆかねばならないと思われる。そこには、運動系、体制感覚系、学習系といった様々な脳の高次機能が関わっている。一方で、キー操作の操作性を詳細かつ正確に評価してゆくためには、新しい操作性評価指標を開発し、それらの指標が先の“心理学的・運動生理学的モデル”のどの部分の操作性を反映しているのか、また、各指標がどのように関係することで、キーの操作性が評価されているのか等について明らかにしてゆかねばならない。キー操作に関わる基礎研究と応用研究のさらなる融合が望まれていると言える。

6 おわりに

本論では、“文章入力課題”の遂行に関わる PC キーボードのキー操作の機能モデル(“階層的タスクモデル”)を提示し、そのモデルに基づき、新しい操作性評価モデル(“階層的操作性評価モデル”)を提案した。本論で提案された“階層的操作性評価モデル”の各指標の総合的パターンを数値的に解析してゆくことにより、ユーザの機器操作の総合評価をある程度予測できる操作性評価モデルの構築を目指している。

参考文献

- ・甘利俊一(2007). 脳は理論でわかるか - 学習、記憶、認知のしくみ.『脳研究の最前線』(第12章)下巻(脳の疾患と数理)(理化学研究所/脳科学総合研究センター編 講談社ブルーバックス), pp299-364.
- ・Bernstein, N.(1967). The co-ordination and Recognition of Movement. Oxford, UK. Pergamon.
- ・d'Avella, A., Tresch, M. C. (2002). Modularity in the motor system : decomposition of time-varying synergies. Advanced Neural Information Processing System, Vol.14, pp.141-148.
- ・d'Avella, A., Bizzi, E. (2005). Shared and specific muscle synergies in natural motor behaviors. Proceedings of National Academy of Science. USA. Vol.102, pp.3076-3081.
- ・d'Avella, A., Proton, A., Fernandez, L., Lacquanti, F. (2006). Control of fast-reaching movements of muscle synergy combinations. Journal of Neuroscience, Vol.26, pp.7791-7810.
- ・入来篤史(2011). 運動の仕組み.『脳科学の教科書 神経編』(第4章). 理化学研究所・脳科学総合研究センター編 岩波ジュニア新書.
- ・中原裕之(2007). 快楽が脳を創る.『脳研究の最前線』(第11章)下巻(脳の疾患と数理)(理化学研究所/脳科学総合研究センター編 講談社ブルーバックス), pp233-297.
- ・Oohashi, T., Nishina, E., Honda, M., Yonekura, Y., Fuwamoto, Y., Kawai, N., Maekawa, T., Nakamura, S., Fukuyama, H., Shibasaki, H.(2000). Inaudible high-frequency sounds affect brain activity : Hypersonic effect. The Journal of Neurophysiology, Vol.83, No.6, pp.3548-3558.
- ・清水信行(2009). 入力機器の操作性に関する研究. アルプス電気株式会社 平成20年度委託研究報告書(2009,3,31). いわき明星大学・入力機器の操作性研究グループ(研究代表者：清水 信行)(未刊行).
- ・田村 博(1998).『ヒューマンインタフェース』(田村 博編 オーム社).

- ・富田 新・大原貴弘(2015). PCキーボードのキースイッチの操作性評価に関する研究 – タッチタイピングによる打鍵法の差異がもたらす操作性評価の相違について –. 応用心理学研究, vol.41, No.1, pp.87-97.
- ・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(2002). 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, Vol.10, No.2, pp.103-112.

注

- 1) 本論では、“心理学的・運動生理学的モデル”という用語を一貫して用いている。この用語には、課題遂行において“効果器”のみならず“感覚器”の働きが含まれているということが含意されている。例えば、“文章入力課題”の遂行においては、“視覚ディスプレイによる入力結果の表示”や“表示内容の視知覚の働き”が不可欠になる。このため、“視覚”を中心とした感覚機能の関与が想定される。このように、“効果器”を中心とした“遠心路”の機能のみならず、“感覚”や“知覚”等の“求心路”の機能が関与することを想定し、また、“文章入力課題”の遂行には、より高次レベルの“心理的機能”が関わっていることも考慮して、本論では“心理学的・運動生理学的モデル”と記述することにした。
- 2) 本論で提示される“階層的操作性評価モデル”は、現時点では、具体性・実用性のあるモデルではない。本論で提示されるモデルは、各機器の操作や課題に対応させて、より具体的・実用的な操作性評価モデルを構築してゆくための概念図のようなものである。このような概念図を“モデル”と呼ぶことには異論があるかもしれないが、本論では“階層的操作性評価モデル”と表現することとした。
- 3) 「入力機器の操作性に関する研究」(アルプス電気株式会社委託研究：平成17年度～平成19年度：研究代表者：いわき明星大学・科学技術学部・システムデザイン工学科・教授 清水信行；研究協力者：同大学・人文学部・心理学科・准教授 富田 新；同学科・准教授 大原貴弘；同大学・理工学部・機械工学科・修士課程2年 東 哲也；肩書きは当時のもの)。
- 4) 富田 新 (2014) PCキーボードの新しいユーザビリティ評価モデルの提案. 応用心理学第81回大会 (於：中京大学) 発表論文集, p17.
- 5) もちろんこれら3階層以外の階層を想定することも可能である。実際、図1に示した階層的タスクモデルは4階層となっている。想定すべき階層数は、課題遂行において使用される機器の種類、ユーザの機器操作法、その機器操作法を心理学的・運動生理学的観点から眺めたときどのように分解するのが最も妥当か等の諸条件によって異なってくる。よって、正確には、“課題遂行の心理学的・運動生理学機能に対応した複数の異なる階層(レベル)を想定し”と記述すべきかもしれない。
- 6) 各階層には異なる脳や中枢神経の部位が関与していると考えられている。例えば、“構想”のレベルには大脳皮質の連合野(運動前野)・小脳・大脳基底核などが、“計画”のレベルには一次運動野・小脳などが、また、“実行”のレベルには脊髄の運動ニューロンなどが関与していると考えられている。
- 7) この事実は本プロジェクトにおいても確認された。例えば、ストロークのない特殊なキーボードを操作するように教示された被験者の筋電パターンは、ストロークのないキーボードへの順応・学習によって変化してゆることが確認された。筋電の測定及び分析は清水信行研究室の修士課程2年 東 哲也(当時)によって行われた(清水, 2009)。
- 8) ここでは“クリック”という言葉を、“機器にセットされたONに関するフィードバック信号”として用いている。つまり、機器設計の段階であらかじめセットされた物理的パラメータのことを指す。これに対して、機器に装備されたこの種の信号を感知し、ユーザの内部で喚起される心理的・生理的印象(例えば、“ON”(キー押しの目的が完了したこと)の印象)は“クリック印象”または“クリック感”と表現される。
- 9) この可能性については本プロジェクトにおける事象関連電位(P300)の測定実験の結果からも示唆されている。複数の被験者のP300のデータが、キースイッチにセットされた“クリック”(“ドロップ荷重”)を感知している可能性を示唆していた。
- 10) 筋電の測定と分析は、理工学部・機械工学科・清水信行研究室(当時)の東 哲也が担当した。

- 11) 評価指標間の相関分析は、人文学部・心理学科・大原貴弘准教授（当時）によって行われた。
- 12) 図2の評価指標の割り付けはあくまでも暫定的なものに過ぎないと言える。モデルに投入する操作性指標の妥当性を検証する作業を今後行ってゆかねばならない。各指標のできるだけ正確な測定を行い、そこで得られたパラメータの関連性を数値的に解析する。このような作業を繰り返すことによって初めて、実用的で信頼性の高い“階層的操作性評価モデル”が得られることになる。
- 13) 例えば各パラメータ間の因果連鎖モデルを考案し、共分散構造分析を用いることによって構造化モデリングを行うという方法は、有力なモデル構築法の1つであろう。

(とみだ あらた／実験心理学)