

ユーザビリティ評価法の評価の論点（信頼性・妥当性・完備性・大規模化・解決案評価）とウェブ認知ウォークスルー（CWW）

Issues on Evaluation of Usability Evaluation Methods – Reliability, Validity, Completeness, Scalability, Redesign – and Cognitive Walkthrough for the Web

北島 宗雄

2006年1月4日

従来の代表的なユーザビリティ評価法は信頼性・妥当性・完備性など、評価法が備えるべき要件を十分に満たしていないという指摘がある [11, 12, 13]。本解説では、これらの指摘について紹介するとともに、それらを克服するユーザビリティ評価法として我々が開発しているウェブ認知ウォークスルー（CWW：Cognitive Walkthrough for the Web） [6, 4, 5] を紹介する。

1 従来のユーザビリティ評価法の問題点

Gray と Salzman [11, 12] は、1990年代に実施された非常に大きな影響を及ぼしたユーザビリティ評価に関する実験的研究を、方法論の観点から批判している。例えば、実験計画や統計解析が適切に行われていないといった問題である。

これらの研究では、ユーザビリティ工学の手法がとられている。具体的には、以下の項目にしたがって、デザイン行為が系統的かつ実際に実施されている：

- ユーザインタフェースデザイン・評価を行うための要求仕様を定める、

- ユーザビリティ問題を突き止める、
- それを解決するデザイン解を創出する、そして、
- その妥当性を検証する。

しかし、Gray と Salzman [11, 12] は、まさにこれらの“「ユーザビリティ工学」分野における研究の基本的な考え方”、つまり、“ユーザインタフェースのデザイン・評価を、秩序や規律が要求される「工学的」なアプローチにより行う”という考え方が、これらのユーザビリティ評価に関する研究の不備の原因になっていると指摘している。

Gray と Salzman [11, 12] は、ユーザビリティ研究の弱点として以下の4点を挙げている：

1. 定義：ユーザビリティの概念が明確に定義されていない。
2. 信頼性：複数の評価者がユーザビリティ評価を行ったとき、発見される問題にばらつきがある。また、発見された問題の重大度の評価が一致しない。これらのばらつきは、評価者間ばかりでなく、評価法間でも生じている。

3. 妥当性と完備性の問題: ユーザビリティ評価法を適用して発見された問題が本当に問題なのかどうか (妥当性) また、発見されたユーザビリティ問題で全てが尽くされているかどうか (完備性) が明らかでない。ここで、妥当性と完備性は、ユーザビリティ評価対象となる局面の全てについて、以下の評価を行うことにより確認することができる:

- hit: 問題があると判定された局面に、実際に問題があった場合の数
- false alarm: 問題があると判定された局面に、実際は問題がなかった場合の数
- miss: 問題がないと判定された局面に、実際は問題があった場合の数
- correct rejection: 問題がないと判定された局面に、実際に問題がなかった場合の数

しかし、これらが適切に行われていない。

4. 解決策の評価: 発見されたユーザビリティ問題に対しては、ユーザビリティ評価結果を踏まえた解決策を考案し、それに基づいて再デザインを行い、評価を行うことが必要である。しかし、それらが適切に行われていない。

2 評定者効果: 認知ウォークスルー法における問題の源泉

Hertzum と Jacobsen [13] は、認知ウォークスルー法、ヒューリスティック評価法、発話プロトコル法の3種類のユーザビリティ評価法について、信頼性、妥当性、完備性の観点から検討を加え、それぞれの手法の抱える問題 (これを、彼らは“評価者効果”と呼んだ) の根本原因を明らかにしようと試みた。Hertzum らは、Gray ら [11, 12] が取り上げた研究に他の代表的なものも加えてユーザビリティ評価実験データを再度分析し、Gray ら [11, 12] の指摘を追認している。

本解説では、Hertzum と Jacobsen [13] の行った分析のうち、認知ウォークスルー (Cognitive Walkthrough, CW と略記) [29] に関する部分について紹介する。そして、続く章で、それらの問題を取り除

いたユーザビリティ評価法として、ウェブ認知ウォークスルー (Cognitive Walkthrough for the Web, CWW と略記) [30, 6] を説明する。これは、CW を発展させてウェブサイトユーザビリティ評価を行えるようにしたものである。

2.1 タスクシナリオの設定・ゴール構造の分析が不適切

第一番目の問題点は、ゴール構造の分析が適切に行われていないという点である。Hertzum ら [13] が調べた実験のうち、評価者が同一のタスクシナリオに基づいて評価を実施していたものは3分の1であり、残りの3分の2では、タスクシナリオが利用されていないか、評価者の選択に任されていた。例えば、Jacobsen と John [14] においては、二人の評価者のうちの一人が発見したユーザビリティ問題の43%は、もう一人の評価者には発見されていなかった。これらの評価者は独自にタスクシナリオを設定してCW分析を行ったが、評価対象としたタスクが異なっていたために評価結果に差が生じたものと考えられる。

CWではタスクシナリオに沿ってユーザビリティ問題の発見が行われる。したがって、評価者効果を取り除くためには、同一のタスクシナリオを使用することが必要になる。しかし、CWではタスクシナリオの設定指針は提供されていないので、評価者は独自にタスクシナリオを設定することになる。一方、タスクシナリオを統一化するだけでは評価者効果を取り除けないことが報告されている [13]。これは、タスクに含まれるゴールがどのような構造を持っているかを明らかにする“ゴール構造分析”が適切に行われないことに原因がある。仮定されるゴール構造が多様であれば、ユーザビリティ評価結果は自ずと多様になってしまう。

2.2 ユーザの認知モデルが分かりにくい

第二番目の問題点は、CWはユーザの操作選択過程について非常に厳密な認知モデル (Polson と Lewis [28] による探査学習のモデル) を仮定しているが、それが評価者に正しく理解されていないという点である。CWは、評価者にユーザ像を想定することを求める。例えば、“Windowsの経験のあまりないユーザの典型例”

をユーザ像として想定することを求める。しかし、この記述はあまりにも漠然としているので、この記述から CW が想定している“厳密な”ユーザの認知モデルに対応するものが構築される保証はない。実際、評価者は、評価者自身に非常に類似したユーザ像を構築してしまう傾向にある [13]。これは、評価者はデザイン対象のドメインに関して非常に深い知識を有している場合が多いが、そのような場合に、自分自身とは異なる知識状態をもつユーザ像を自分自身の知識を用いて想定することは原理的にできないためである。

2.3 ユーザビリティ問題の同定の基準があいまい

第三番目の問題点は、ユーザビリティ評価法の定めるユーザビリティ問題の同定基準が明確でなく漠然としているという点である。これは、評価結果の信頼性と妥当性に大きく影響する。この点に関し、CW では、これまで、ユーザビリティ問題の同定基準を明確にすることに注力してきており、最新のバージョンでは旧来バージョンで用いられていた複雑な評価基準を4つの質問にまとめている [29]。しかし、それでもなお、これらの基準を適用するのは非常に煩雑であり、時間がかかるということが指摘されている [13]。その結果、評価結果の信頼性と妥当性の問題を十分にクリアできていなかった。

2.4 大規模な適用が難しい

第四番目の問題点は、ユーザビリティ評価法を大規模で複雑なアプリケーションに適用することが難しいという点である。ユーザビリティ評価の完備性を達成するためには、評価対象のアプリケーションや機器で達成することが想定されている“全ての”タスクについてユーザビリティの問題があるかどうかを検査されなければならない。ユーザビリティテストは非常に時間がかかるので、実用的な時間で評価結果を出すためには、評価者は評価対象アプリケーション・機器の機能の中から評価対象の機能を選択することを余儀なくされる。このプロセスに標準的なものは存在しない。評価対象からもれた機能が、重大なユーザビリティ問題の源泉になってしまう可能性は原理的に排除

不可能である。

3 ウェブ認知ウォークスルー (CWW)

3.1 CWW の理論的基盤

CWW は、認知モデルに基づくユーザビリティインスペクション法 [25] であり、情報を求めてあるサイトを訪れたユーザが問題なくそのサイトで情報にたどり着けるかどうかということについて、ユーザビリティ問題の同定、改善法の提案を行う [6, 4, 5]。CWW は、ユーザのウェブサイトナビゲーションプロセスをシミュレートする認知モデル CoLiDeS モデル [17] に基礎を置いている。これは、CW が探査学習のモデル [28] にユーザの操作選択プロセスのシミュレーションの基礎を置いているのに対応する。

CoLiDeS モデルは、Comprehension-based Linked model of Deliberate Search (意図的探索の理解に基づくネットワークモデル)の略であり、デスクトップアプリケーションにおいて正しい操作を探索により発見し実行するプロセスをシミュレートする認知モデル (LI-CAI モデル [19]、Linked model of Comprehension-based Action planning and Instruction taking の略)をウェブ探索に適合するように拡張したものである。これらは、いずれも文章理解・問題解決の認知モデルである Kintsch [16] の構築統合理論 (Construction-Integration Theory) にその基礎を置いている。

CoLiDeS モデルでは、ユーザの情報探索プロセスとして次のものを仮定する：“ユーザは、訪問したウェブサイトにおいて獲得したい情報を求めてサイト内ページのナビゲーションを行う。その際に、ユーザは探索ゴールに最も類似したアクションを選択することによる前方探索を行う”。ここで、“アクション”は、心的行為、運動行為の両方を意味し、前者には、ウェブページ内の特定の領域(例えば、サイトナビゲーションバー、サイトロゴ、プロモーションなど)に注意を向ける、などが、後者には、リンクをクリックする、などが含まれる。この仮定は、ウェブサイトユーザビリティに関する他の研究(例えば、[8, 7, 9, 10, 15, 23, 24, 27])と共通した仮定である。これらの研究では、類似の程度を示す言葉として、“情報香 (information scent) [26, 10, 8, 7] ”

が広く用いられている。

図 1 は、CoLiDeS モデルによるユーザのリンク選択プロセスのシミュレーションを模式的に示している。ここで想定している状況は、「あるユーザが、「休暇を利用して家族でニュージーランドに旅行し、ハイキングをしようと考えている。…」という情報探索ゴールを抱いてこのサイトを訪れた」というものである。シミュレーションは以下のように進む。

ユーザは、まず、ページの全体をパージングし、7つの部分領域に分割する。そして、コンテンツ領域を、次に操作を行うべき領域として選択する。さらに、選択した領域をパージングし、次に注意を向けるべき領域を選択しようとする。しかし、一つに絞り込むことができない。最左端の領域 International と Other Sites が候補に挙がるが、前者を選択する。次に、この領域に含まれるリンク、Africa、Asia、Europe、North America、Oceania、South America、Intl. Park/Govt. から一つを選択しようとする。しかし、いずれのリンクもゴールとの関連が深いとは判断できず、どれも選択できない。

ここで、正しいリンクは Oceania であったが、上記シミュレーションでは、ユーザの持っている知識では New Zealand と Oceania を関連付けることができなかった。つまり、このユーザは、Oceania から十分な情報香を嗅ぎ取れなかった。後述するが、CWW では、想定したユーザの知識に対してリンクが放つ情報香の定量的評価を行い、このような問題の発見を行っている。

3.2 CWW が同定するユーザビリティ問題

前節で説明したように、CoLiDeS モデルは、ある情報を獲得しようとしてあるサイトを訪れたユーザが前方探索を行ってサイト内ナビゲーションを行うプロセスをシミュレートする。CWW は、このプロセスがうまくいくかどうかという観点から、サイト内ページのユーザビリティ評価を行う。図 1 では、CoLiDeS モデルによるシミュレーションに対応させて、右側に CWW によって発見されるユーザビリティ問題を示した。以下に、個々のユーザビリティ問題について説明する：

1. 情報香の弱いリンク (weak scent link): 正しいリンクとユーザゴールの意味的な類似性の程度が弱いという問題。このような場合、ユーザは正しいリンクの意味は適切に理解するものの、それが見込みのある操作対象であると判断する可能性が低い。
2. 馴染みのない見出し・リンク (unfamiliar heading/link): 正しいリンクやそれが含まれている領域の見出しを理解するのに必要な知識を欠いているという問題。リンクや見出しに出現頻度が低い言葉が使われていたとき、たとえ、探索ゴールとの意味的類似性が高かったとしても、それを選択するに十分な情報香が知覚されない。例えば、図 1 の Oceania と New Zealand は、意味的には類似しているが Oceania の出現頻度が低いので情報香は小さい。
3. 競合する見出し (competing headings): ある見出しによって指定される領域について、その領域内のリンクを選択しても探索ゴールを達成できないにもかかわらず、その領域が十分な情報香を発しているという問題。この問題は、重大な問題に繋がる危険性がある。CoLiDeS モデルでは、ユーザは、競合する見出しに独占的に注意を向け、その領域内のリンクのみを選択対象とすると仮定している。その結果、正しいリンクを含む領域には注意が向かない。間違ったところにはまったらかななか抜け出せないというガーデンパス問題が生じる。
4. 競合するリンク (competing links): 正しい領域、もしくは、競合する領域に、探索ゴールに意味的に類似しているが正解ページには導かないリンクが存在するという問題。

3.3 潜在意味解析 (LSA) と情報香

CWW では、探索ゴールと見出し・リンクとの意味的類似度を潜在意味解析 (LSA: Latent Semantic Analysis [21, 20, 22]) を用いて計算する。以下に、LSA について、解説記事 [31] を引用し、一部補足しながら説明する：

潜在意味解析とは、さまざまな文脈におい

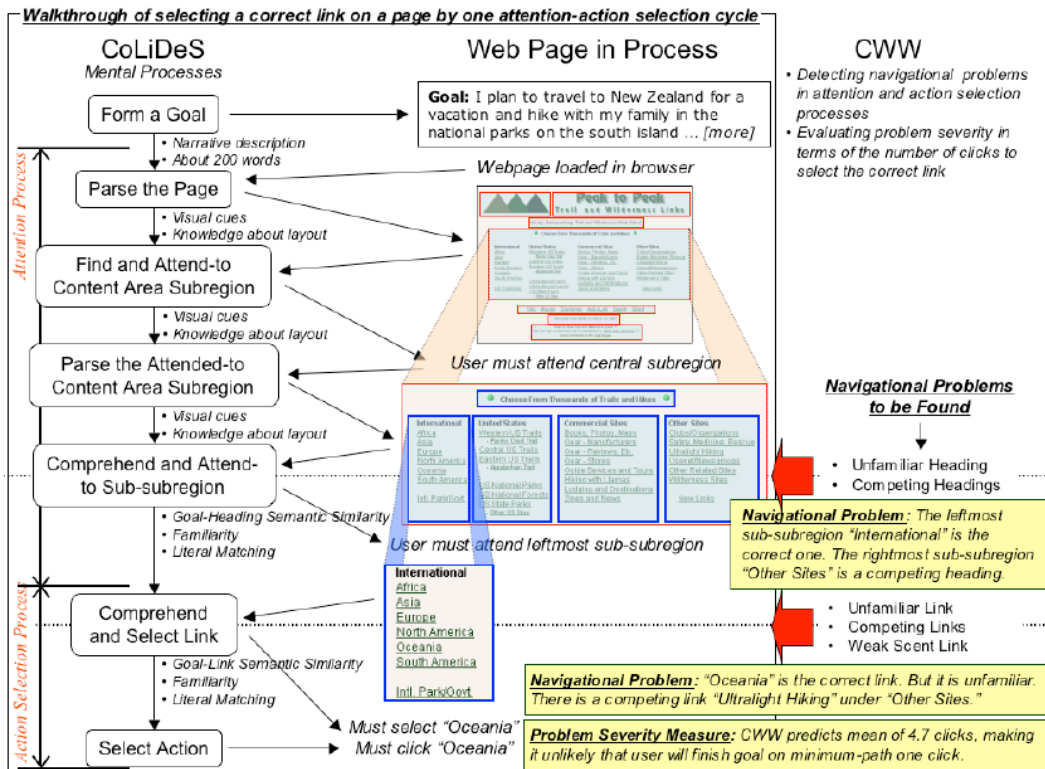


図 1: CoLiDeS model for how user accomplishes goal on Peak to Peak home page (<http://www.peaktotoppeak.net>) with CWW problem identification and prediction of 4.7 mean total clicks to accomplish user goal on this webpage. 文献 [5] より転載

て語の意味がどのように使用されているかを大規模なテキストコーパスに現れるすべての語や語の集合(文、節、小論など)に対し統計的な計算(特異値分解)と次元縮約を施すことによって導出し意味空間として表現する理論である。この理論の根幹をなす考え方は、文脈によって特定の語が現れたり現れないということが、語と語、あるいは語の組み合わせの間の意味的な類似性を決定するというものである。意味空間は以下の手続きで構成される。まず、テキストコーパスが与えられたとき、ある語がある文脈で現れる頻度を適当に変換して得られる“語 - 文脈”行列 A (n 語 $\times m$ 文脈, 階数 r) を作成し、 $A = U D_{\lambda} V^T$ の形に分解する(特異値分解)。ここに、 D_{λ} は $\lambda_1 \geq \dots \geq \lambda_r > 0$ (特異値) を対角要素にもつ対角行列、 U 、 V は正規直交ベクトルを列ベ

クトルにもつ行列 ($U^T U = V^T V = I$) である。 $U = (u_1, \dots, u_r)$ 、 $V = (v_1, \dots, v_r)$ とおけば、 $A = \sum_{i=1}^r u_i \lambda_i v_i^T$ と表せる。ここに、 λ_i は特異値、 u_i 、 v_i は左あるいは右特異ベクトルと呼ばれる。次に、行列 A の階数を r から k に減じることにより $A_k = \sum_{i=1}^k u_i \lambda_i v_i^T$ を作成する。 k は通常 50 ~ 1,000 である。この行列は、階数を k とした場合の行列 A の最適近似となることが知られている。それぞれの語や文脈は、左、右特異ベクトルの k 個の要素を用いて k 次元空間のベクトルとして表現される。次元縮約後の行列 A_k においては、もとの行列 A に含まれている語の使用法が多様であることに起因するノイズが除かれ、意味的に近い語が、たとえそれらが同じ文脈に現れていなくても近い場所に配置させられている。米国大学生の言語的知識に対応する意味空間が作成されているが、その場合、

$n = 92,409$, $m = 37,651$, $k = 419$ である。このようにして作成される意味空間が人間の知識を反映していることは、さまざまな方面で検証されている。たとえば、潜在意味解析によって作成された意味空間を用いて語彙テストや読解テストを行ったとき、人間の成績と似たパターンを示すことや、語の分類やカテゴリー判断の結果が人間の行った結果に似ていることが報告されている。

なお、CWW ウェブサイト [1] では、現在、フランス語（大学生レベル）、英語（3年、6年、9年、12年、大学生レベル）の意味空間が提供されている。

このように、潜在意味解析によれば、語や語の集合は高次元意味空間内の一点 \vec{a} として表現されるので、2つの語・語集合 \vec{a} , \vec{b} の間の距離は、それらベクトル間の関係を定量化して与えられる。CWW では、これらの間の距離は $\cos(\vec{a}, \vec{b})$ で与えられる。意味的な類似性が高い場合には1に近い値になる。関連がなければ0に近い値になる。

CWW では、ユーザは探索ゴールとの意味的類似度が最も大きな見出し領域、リンクを選択すると予測する。意味的類似度を評価する際には、探索ゴールとしては“ニュージーランド旅行”というような簡潔な記述ではなく、ユーザが探索時に抱いていると考えられる動機や探している情報の種類などを含む具体的な記述を用いる。こうすることにより、さまざまな文脈で行われる可能性のある“ニュージーランド旅行”の曖昧性を除去し、意味空間内において、適切に当該ユーザの探索ゴールを表現する点を特定する。これと同様の理由により、簡潔に表現されている見出しやリンクも、具体化・精緻化される。手続きの詳細については、[3]を参照のこと。

LSA が提供するもう一つの重要な指標は、ベクトル長である。これは、語の出現頻度と相関があり、その語に対する知識の深さを表している。

3.4 CWW の信頼性・完備性

CWW で同定できる4種類のユーザビリティ問題は、LSA を利用することにより客観的に評価することができる。すなわち、想定されるユーザの意味空間を参照して、探索ゴールと見出し・リンクとの意味的類

似度は、それらの間の余弦として得ることができる。また、馴染みの程度は、語のベクトル長に対応付けて評価することができる。こうすることにより、評価者の主観を全く介入させずにユーザビリティ問題を発見することができる。現在、自動的にユーザビリティ問題を発見するルールとその根拠がウェブに公開されている [2]。また、問題の程度を評価する式が開発されている。その式では、あるページにおいて正しいリンクを選択するのに必要とされるマウスクリック数 N は、以下の式で与えられる [5]：

$$N = 2.199 + 1.656 \times N_1 + 1.464 \times N_2 + 0.754 \times N_3$$

ここで、 N_1 はそのページの正しいリンクに馴染みがない場合、すなわち、1語リンクでベクトル長が0.55未満、2語以上のリンクでベクトル長が0.8未満の場合は1、そうでない場合は0、 N_2 はそのページの正しいリンクの情報香が弱い場合、すなわち、余弦が0.1よりも小さい場合は1、そうでない場合は0、 N_3 は競合する見出しの下に存在する競合するリンクの数である。重大な問題がある場合とは、 $N > 5$ の場合である。

3.5 CWW の妥当性

CWW を用いたユーザビリティ評価実験は、これまでに、数百人の被験者を用いて行われてきている [6, 4]。これらの実験では、Microsoft Encarta 百科事典を模倣した実験用サイトにおいて、被験者が探索をどのように行うかが観察された。被験者には、探索すべき記事の概要が200語程度で提示される。被験者は、カテゴリー毎に分類された約90のトピックから、概要に示された記事に到達できると考えられるトピックを選択する。CWW の言葉に対応付ければ、カテゴリーが見出しに、トピックがリンクに対応する。これらの評価実験結果を検討することによって、CWW の妥当性を検討することができる。すなわち、(1) hit、(2) false alarm、(3) miss、(4) correct rejection の4つの指標を用いて、CWW が正しくユーザビリティ問題を発見している程度を評価できる [5]。

ユーザビリティ問題の重大さを、高 ($5.0 < \text{クリック数}$)、低 ($2.5 < \text{クリック数} \leq 1.0$)、なし ($1.0 \leq \text{クリック数} \leq 2.5$) に分類し、64のタスクについて検討したところ、問題が高または中と判定された29タスクについて、(1) hit 率90%、(2) false alarm 率10%であった。また、問題がないと判定された34タスクにつ

いて、(3) miss 率 31%、(4) correct rejection 率 69%であった。miss 率がやや高いが、見逃されたものの中で重大度が高いものはなかった。

3.6 解決案の評価

CWW を実施すれば問題の性質に応じて修復したページを作成できる。例えば、競合するリンクを介してゴールが達成できるようにすれば、その問題は解決される。また、馴染みのない語が使われている場合は、意味的に似ている馴染みのある語に置き換えればよい。Blackmon ら [5] は、一連の実験 [6, 4] においてユーザビリティ問題が生じないように修復したページを作成し、それらが確かに問題を生じさせず修復の効果がかったことを、前節と同様の指標により評価している。

3.7 CWW の大規模ユーザビリティ評価への適用

Kitajima ら [18] は、Microsoft Encarta のウェブサイトの全ての記事ページ (30,000 ページ以上) について、そのページの概要が書かれていると考えられる第一段落を探索ゴールとしたときの、ユーザのカテゴリ選択、トピック選択プロセスのシミュレーションを、トピック・リンク選択プロセスを CoLiDeS モデルの特徴を取り入れた Markov 連鎖によりモデル化することによって行った。また、ユーザビリティに問題のある場合、すなわち、与えられたゴール記述を用いた場合に正しいカテゴリ、正しいリンクを選択できない場合について、ユーザビリティ問題の性質に応じて修復を施して構築される“修復済みサイト”についても同様のシミュレーション実験を行い、修復の効果を評価した。ここでは、いくつかの修復法の改善度効果の評価を行っている。ウェブサイトにおいては、訪問者が求めている情報に途中のナビゲーションページを介して確実にたどり着けるようにすることが必要であるが、それをウェブサイトが満たしているかどうかを評価し、問題のある箇所を修復したときの効果を事前に見積もるといった方法の可能性を探るのがこの研究の狙いである。

このような、大規模ユーザビリティ評価は、CWW が LSA を用いていること、ならびに、客観的で人手

によらないユーザビリティ問題発見法を定めていることにより可能となった。

4 まとめ

ユーザビリティ評価法の問題として挙げられている信頼性、妥当性、完備性、解決案評価、大規模化の問題に対し、CWW では、次のように応えている。まず、信頼性、完備性については、従来、評価者の主観に依存していた部分に LSA を導入することにより評価の客観化を実現している。ヒューリスティック評価法やインスペクション法を実施する場合、評価者はユーザに成り代わって評価することが求められる。しかし、これは原理的に不可能なことであり、たとえ、ユーザビリティ評価のエキスパートといえども、本人が気づかないうちに、評価者自身の個人的な知識を反映した評価を行ってしまっているという可能性は否定できない [13]。CWW では、評価ルールを定義することにより、評価者の主観が入り込む余地をなくしている。

妥当性については、一連の実験結果 [6, 4, 5] をもとに実証している。Gray と Salzman [11, 12] は、これまでのユーザビリティ評価の研究が妥当性を示すためのデータを提出してきていないことを批判している。また、ユーザビリティ問題の重大度を示すこともできていないと指摘している。これらの点について、CWW では、問題の重大度を計算する式を示し問題への対処方針を決定するのに必要な情報の提供を行い、これらの問題を解決している。

解決案についても、CWW は問題の所在を明確に示すことができるので、問題の性質に応じた解決案を提示できる。そして、解決案を反映した再デザインも、旧デザインと同様に客観的に評価される。最後に、大規模化についても、CWW 手続きの自動化により対処できる。

現在、CWW は、世界的に関心を集めており、その核となるそれぞれの言語、ユーザ層に対応した LSA の構築が進められている。また、簡便に CWW を行う為の環境も整備されつつある。信頼性、妥当性、完備性、大規模適用可能性を備えたユーザビリティ評価法として、今後、広く利用されるようになることを期待している。

参考文献

- [1] <http://autocww.colorado.edu/>.
- [2] <http://autocww.colorado.edu/blackmon/Tutorials/AutomatableRules.doc>.
- [3] Blackmon, M., Kitajima, M., Mandalia, D. and Polson, P.: *LSA: A Road to Meaning*, Erlbaum, Mahwah:NJ, 2006, chapter Automating Usability Evaluation: Cognitive Walkthrough for the Web Puts LSA to Work on Real-World HCI Design Problems.
- [4] Blackmon, M., Kitajima, M. and Polson, P.: Repairing Usability Problems Identified by the Cognitive Walkthrough for the Web, in *Human Factors in Computing Systems (CHI2003)*, 2003.
- [5] Blackmon, M., Kitajima, M. and Polson, P.: Tool for Accurately Predicting Website Navigation Problems, Non-Problems, Problem Severity, and Effectiveness of Repairs, in *Human Factors in Computing Systems (CHI2005)*, 2005.
- [6] Blackmon, M., Polson, P., Kitajima, M. and Lewis, C.: Cognitive Walkthrough for the Web, in *Human Factors in Computing Systems (CHI2002)*, 2002.
- [7] Chi, E., Pirolli, P., Chen, K. and Pitkow, J.: Using information scent to model user information needs and actions and the Web, in *Human Factors in Computing Systems (CHI2001)*, 2001.
- [8] Chi, E., Pirolli, P. and Pitkow, J.: The scent of a site: A system for analyzing and predicting information scent, usage, and usability of a website, in *Human Factors in Computing Systems (CHI2000)*, 2000.
- [9] Chi, E., Rosien, A., Supattanasiri, G., Williams, A., Royer, C., Chow, C., Robles, E., Dalal, B., Chen, J. and Cousins, S.: The Bloodhound Project: Automating discovery of web usability issues using the InfoScent? Simulator, in *Human Factors in Computing Systems (CHI2003)*, 2003.
- [10] Furnas, G.: Effective view navigation, in *Human Factors in Computing Systems (CHI1997)*, 1997.
- [11] Gray, W. and Salzman, M.: Damaged merchandise? A review of experiments that compare usability evaluation methods, *Human-Computer Interaction*, Vol. 13 (1998), 203–261.
- [12] Gray, W. and Salzman, M.: Repairing damaged merchandise: A rejoinder, *Human-Computer Interaction*, Vol. 13 (1998), 325–335.
- [13] Hertzum, M. and Jacobsen, N.: The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods, *International Journal of Human-Computer Interaction*, Vol. 15 (2003), 183–204.
- [14] Jacobsen, N. and John, B.: Two case studies in using cognitive walkthrough for interface evaluation, Technical Report CMU-CS-00-132), Carnegie Mellon University, Pittsburgh, PA, 2000.
- [15] Katz, M. and Byrne, M.: Effects of scent and breadth on use of site-specific search on e-commerce websites, *ACM Transactions on Computer-Human Interaction*, Vol. 10 (2003), 198–220.
- [16] Kintsch, W.: *Comprehension: A Paradigm for Cognition*, Cambridge University Press, 1998.
- [17] Kitajima, M., Blackmon, M. and Polson, P.: A Comprehension-based model of Web navigation and its application to Web usability analysis, in *People and Computers XIV*, Springer, 2000.
- [18] Kitajima, M., Kariya, N., Takagi, H. and Zhang, Y.: Evaluation of Website Usability Using Markov Chains and Latent Semantic Analysis, *IEICE TRANS. COMMUN*, Vol. E88-B (2005), 1467–1475.
- [19] Kitajima, M. and Polson, P.: A comprehension-based model of exploration, *Human-Computer Interaction*, Vol. 12 (1997), 345–389.

- [20] Landauer, T.: Learning and representing verbal meaning: Latent Semantic Analysis theory, *Current Directions in Psychological Science*, Vol. 7 (1998), 161–164.
- [21] Landauer, T. and Dumais, S.: A solution to Plato’s problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge, *Psychological Review*, Vol. 104 (1997), 21–240.
- [22] Landauer, T., Foltz, P. and Laham, D.: An introduction to Latent Semantic Analysis, *Discourse Processes*, Vol. 25 (1998), 259–284.
- [23] Larson, K. and Czerwinski, M.: Webpage design: Implications of memory, structure and scent for information retrieval, in *Human Factors in Computing Systems (CHI1998)*, 1998.
- [24] Miller, C. and Remington, R.: Modeling Information Navigation: Implications for Information Architecture, *Human-Computer Interaction*, Vol. 19 (2004), 225–271.
- [25] Nielsen, J. and Mack, R.: *Usability Inspection Methods*, John Wiley and Sons, New York, 1994.
- [26] Pirolli, P.: Computational models of information scent-following in a very large browsable text collection, in *Human Factors in Computing Systems (CHI1997)*, 1997.
- [27] Pirolli, P. and Card, S.: Information foraging, *Psychological Review*, Vol. 106 (1999), 643–675.
- [28] Polson, P. and Lewis, C.: Theory-Based Design for Easily Learned Interfaces, *Human-Computer Interaction*, Vol. 5 (1990), 191–220.
- [29] Wharton, C., Rieman, J., Lewis, C. and Polson, P.: *Usability Inspection Methods*, John Wiley, New York, 1994, The cognitive walkthrough method: A practitioner’s guide, 105–140.
- [30] 北島宗雄：ウェブ認知ウォークスルーによるウェブサイトユーザビリティの評価, 日本ファジィ学会誌, Vol. 14 (2002), 446–460.
- [31] 北島宗雄：用語解説「潜在意味解析 (Latent Semantic Analysis: LSA)」, 日本知能情報ファジィ学会誌, Vol. 17 (2005), 76.