

定量的ユーザビリティ評価手法： NEMによる操作性の評価事例およびツール開発の報告

鱗原晴彦*¹ 龍淵 信*¹ 佐藤大輔*¹ 古田一義*²

A case study of "efficiency of an operation" evaluation and the following tool development by NEM(quantitative usability evaluation method)

Haruhiko Urokohara *¹, Makoto Tatsubuchi *¹, Daisuke Sato *¹ and Kazuyoshi Furuta *²

Abstract - On developing a new product, we acknowledge that staged measurements of "efficiency of its operation" turn out to be quite important. The NEM usability evaluation support tool, which we have developed, aims at its importance and we now completed the development of a data analysis system after operation time is measured. Its original objective is to detect a gap between a novice(a general user) and an expert(a designer) models. Going a step further, we have also achieved to examine a gap between variously ranked user models without difficulty or loss of time. We report here the evaluation method with visualized "efficiency of an operation" and its effectiveness.

Keywords: Usability evaluation, Task performance measure, User Segment, Novice Expert ratio Method, NEM

1. はじめに

NEM(Novice Expert ratio Method)^[1]は操作時間を比較することによって、ユーザのモデルとデザイナーのモデルとのギャップ (NE比) を抽出しようとしたものである。1999年に本格的な考察をはじめてから現在までに10例について評価を実施し、商品開発の現場へのフィードバックをおこなってきた。

操作時間を測りNE比を分析するだけで問題個所の発見ができることを目指したNEMだが、事例が増えるにつれてNE比以外の様々なギャップについても考察の必要性が生じてきている。しかし、これらの多様な分析を手作業で行っているのは、評価業務をより早くより正確に実施するためにも限界があるため、専用ツールの開発に着手した。本発表においてはNEMの要にあたる操作時間比の分析ツールについてその効果と今後の可能性について報告する。

2. NEMの概要

ノーマンのメンタルモデル(Three Aspects of Mental Models:1986)ではデザイナーの描くデザインモデルとユーザの描くユーザモデルとに生じるギャップが、使いにくさの要因になると提唱している。このモデルをそのまま定量的にビジュアライズしようと試みたのがNEMである。すなわち、商品開発を担当し操作性に熟達したデザイナーの操作時間に対し、一般ユーザもしくは初心者ユーザが要する操作時間との比をNE比とし、この比によってギャップを可視化している (図1)。

タスクをいくつかの操作ステップ (サブタスク) にわけてNE比を分析してみると、問題点の内在するサブタスクを抽出することが可能となる。

また、操作ステップの区切り方1つで、特定のフェー

ズの認知処理に要する時間の比較や、作業部分に要する時間のみの比較など、様々な分析アプローチが可能である。

さて、一般ユーザの操作時間が長く、かつデザイナーの操作時間も長ければNE比は小さくなる。この場合の操作タスクでは両者のモデルが一致している可能性が高い。この場合、操作性の見直しをするにはデータベースの改善や運用の改善など実現に多大な工数を要する場合が多く、容易に改善することは難しい。

一方、NE比の高い操作タスクには両者間のモデルギャップが内在すると考えられる。NE比が高い操作タスクは、デザイナーがユーザの気持ちを理解しさえすればNE比を小さくする、すなわち、そのギャップを小さくすることが可能で、操作性を改善しやすいタスクといえる。このことは現状の設計条件、運用条件下において適正な工数で改善できるタスクであることを示唆しており、またNE比そのものが操作性の改善項目のプライオリティを判断する指標としても活用できる。

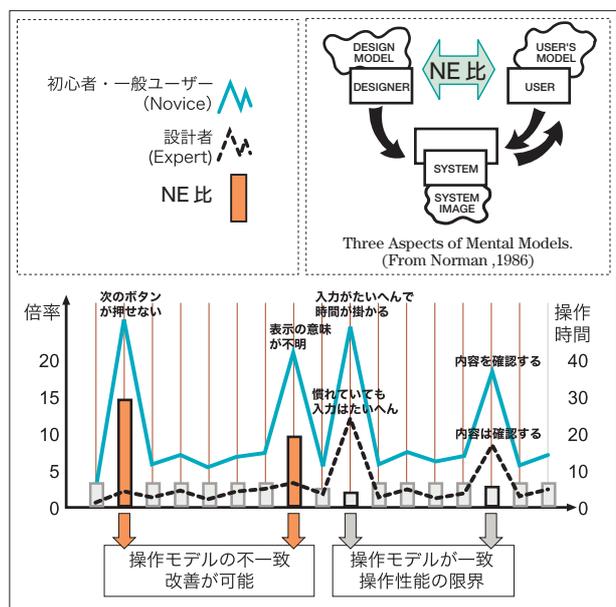


図1 NEMについて
Fig1. NEM outline

*1: 株式会社ユー・アイズ・ノーバス

*2: 株式会社ユー・アイズ・ノーバス 使いやすさ研究所

*1: U'eyes novas Inc.

*2: U'eyes novas Tsukaken Lab.

3. NEMによる評価事例

弊社におけるUCD活動の中で、以下のシステムや商品に対してNEM評価を実施し、問題点の発見およびその改善提案に役立てた。この評価業務から得た様々な考察ノウハウを基に分析ツールの要求仕様の検討を行った。

- ・ ATM
- ・ カーナビゲーション (2機種)
- ・ デジタルカメラ
- ・ カーエアコンパネル (3機種)
- ・ 新規操作デバイス
- ・ プリンタードライバー
- ・ カーナビ用スピーチインタフェース

4. NEMの評価手順

NEM評価では、操作時間の抽出作業・操作ログ解析作業・分析を容易にするグラフ化作業の3つ(図2)を手際よく行なう必要がある。特に操作ログの解析作業では、分析を行うステップの区切り方などの試行錯誤が自在にできることが重要となる。

現状の評価作業手順を以下に示す。

- 1) 評価対象システムの操作キーから操作ログを抽出できるような仕組みを構築
 - ・ GUIの場合マウスクリックによるイベントを抽出
 - ・ ハードスイッチ、タッチパネルの場合は直接押下イベントを抽出
 - ・ 赤外線信号、U-ART信号などからの抽出
 - ・ ビデオ観察によるマーキングなど
- 2) 評価目的およびタスク内容によりどの操作時間を抽出すべきか試行錯誤を繰り返し考察する
- 3) グラフ化しNE比の高いタスクについてその原因を考察する
 - ・ 平均値より高いタスク
 - ・ 他の操作タスクと比べて著しく高いタスク
 - ・ グラフに現れるユーザ固有の操作リズムが崩れているタスクなど
- 4) 達成度、操作時間のバラツキ、学習効果などから総合的な判断を行なう

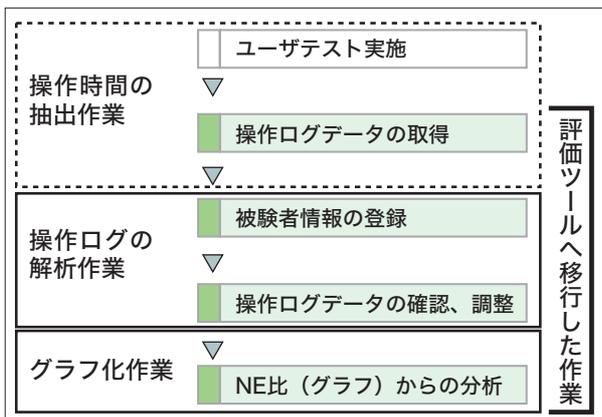


図2 NEM 評価の作業手順
Fig2. NEM evaluation process

5. NEM分析ツールの概要

NEM評価を行う場合、これまではユーザテストから得られた操作ログの解析を手動で行っていた。例えば、あるタスクを解析する時には、被験者が評価機を操作した大量の操作ログデータの中から設定したサブタスクの正解ステップを見つけ出す、という非常に非効率な作業を根気よく続けていた。こうした状況を打開することが本ツール開発の第一の目的である。

5.1 NEM評価業務上の作業効率の改善項目

- 1) 自動で操作ステップの正解判定ができる
 - ・ 本ツールを用いると正解ステップの自動判定ができるため、解析作業効率が格段に向上する。
- 2) 手動でも解析が行える
 - ・ ユーザテストの進行や評価モデル上の不具合、被験者の意図しない状況などによって不自然なログデータが抽出されることがある。解析対象から外したいログデータや正解ステップの修正など、よりの確な解析のための調整作業を手動で任意に素早く行なうことができる。
- 3) NE比のグラフ化が容易にできる
 - ・ NEM評価手法の特長は、操作モデルのギャップをグラフによって可視化することにある。本ツールではユーザ属性や解析条件を変更しながら都度、容易にグラフ化することが可能である。

5.2 NEM分析ツールによる作業手順

- 1) 情報を登録する
 1. 被験者情報の登録
 - ・ エキスパート、ノービス、性別、年齢などの基本的なユーザ属性以外にも当該機種の評価に必要なユーザ属性を登録可能となっている。
 2. テスト機情報の登録
 - ・ テスト機名称や評価するスイッチの種類、動作仕様などを登録する。
 3. 正解判定情報の登録
 - ・ NEM評価に有効なサブタスクを設定し、各サブタスク毎に正解操作となる操作ステップ状態を登録する。
- 2) ログデータを取得する
 1. ログ取得ツールによって、実験中に評価機種から出力された操作ログをツール内に取り込む
- 3) ログデータを解析する
 1. 解析処理をする(図3)
 - ・ 操作ログ情報は、CSV形式から様々な分析考察に活用可能なXML形式へ変換される。サブタスクの判定を自動で行う(手動も可能)。
 2. グラフ化する(図4)
 - ・ 全ユーザーによるNE比のグラフ化から、様々な属性間の比までを適宜グラフ化し分析・考察を行える。

6. ツールがもたらす効果

6.1 分析作業の効率化

一般に、操作時間を計測し抽出データを分析するためには多大な労力を要する。開発プロセスにフィードバックするため、できるだけ短時間に低コストで評価を実施しようとするには難しい手法と言わざるを得ない。特にユーザ像も、提供される機能も、サービスも多様化した情報機器の操作性を検討しようとするといへんな手間がかかる。NEMも例外ではなく操作時間を計測して単純に比較するだけであるにも関わらず、いざ分析しようとする場合に試行錯誤に要する工数は大きい。本ツールの登場により一操作に要するログデータの分析コストは従来の1/5以下となり、その分、問題点の考察に費やす工数を増やすことができるようになった。

6.2 問題タスク発見率の向上と考察力の向上

専門家評価の効果は高く、その作業効率もよいものの、同じ市場における同じ商品を繰り返し評価するような場合には、やはり業界の常識、固定観念が邪魔をして客観的な考察力が薄れる場合もある。本ツールでは様々な比較個所を選定し、多くの観点からモデルのギャップを起こしている問題箇所を抽出することができる。経験則的ではあるが複数の専門家が考察をするのと同等の効果が

あるものと期待している。

また、問題発見のための試行錯誤を短期間に繰り返すことができるため、操作モデルギャップに影響を及ぼす本質的な要因を探り出すことも期待できる。

6.3 様々なユーザ属性の比較から様々なモデルギャップを考察する

通常のユーザテストで思いがけない操作をしてしまうユーザに出会うのと同様、NEMの分析中にもユーザによって極端に操作時間に差が出る操作タスクに出会う場合がある。

そのような場合でも、本ツールによる考察は必ずユーザ情報との関係を確認しながら検証できるため、ギャップの大きいユーザがどんな属性なのかを容易に確認することができる。

操作性の検討を行なう場合に性別や年齢による比較を行いたいという要望が多い。もちろん、これらは操作能力を左右する要因のいくつかではあるが、もっと着目すべき要因は多いと考える。

操作性を考える上で、ユーザがシステムに向かう時の気持ちを無視することはできない。ユーザがシステムに対して積極的になれるのはどのような状況なのか、また、どのようなユーザ属性を持つ人が多機能を使いこなせるのかなど、基本属性以外のユーザ属性にもスポットを当

ステップ	コード	キー名称	取得時間	温度	オーディオ	CDトラック	内気/外気	コンピュータ	INFO	フォロ
00000				18	AM/FM	1	内気	FM多重	車速	OFF
00001		スタートキー	11:00:00.00	18	AM/FM	1	内気	FM多重	車速	OFF
00001	C	AUTO	11:00:01.32	18	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:01.98	18.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:02.02	19	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:02.05	19.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:02.08	20	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:03.23	20.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:03.30	21	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:03.80	21.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:03.87	22	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:04.80	22.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:04.83	23	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:04.87	23.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:04.90	24	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:04.93	24.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:06.08	25	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:06.12	25.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:06.15	26	AM/FM	1	内気	FM多重	車速	OFF
00001	W	温度	11:00:06.92	25.5	AM/FM	1	内気	FM多重	車速	OFF
00001	W	温度	11:00:06.98	25	AM/FM	1	内気	FM多重	車速	OFF
00001	W	温度	11:00:07.02	24.5	AM/FM	1	内気	FM多重	車速	OFF
00001	W	温度	11:00:07.08	24	AM/FM	1	内気	FM多重	車速	OFF
00001	W	温度	11:00:07.18	23.5	AM/FM	1	内気	FM多重	車速	OFF
00001	Q	温度	11:00:07.67	24	AM/FM	1	内気	FM多重	車速	OFF
00001*	Q	温度	11:00:07.73	25	AM/FM	1	内気	FM多重	車速	OFF
00002		スタートキー	11:00:07.73	24.5	AM/FM	1	内気	FM多重	車速	OFF
00002	R	オーディオ	11:00:11.80	24.5	CD/MD	温度=25・正解到達まで:7.73秒	内気	FM多重	車速	OFF
00003		スタートキー	11:00:11.80	24.5	CD/MD	1	内気	FM多重	車速	OFF
00003	W	温度	11:00:13.48	24	CD/MD	1	内気	FM多重	車速	OFF
00004	W	温度	11:00:14.08	23.5	CD/MD	1	内気	FM多重	車速	OFF
00004		スタートキー	11:00:14.08	23.5	CD/MD	1	内気	FM多重	車速	OFF
00004	U	CDトラック	11:00:16.20	23.5	CD/MD	2	内気	FM多重	車速	OFF
00004	U	CDトラック	11:00:17.13	23.5	CD/MD	3	内気	FM多重	車速	OFF
00005		スタートキー	11:00:17.13	23.5	CD/MD	3	内気	FM多重	車速	OFF
00005	O	内気/外気	11:00:20.95	23.5	CD/MD	3	外気	FM多重	車速	OFF
00006		スタートキー	11:00:20.95	23.5	CD/MD	3	外気	FM多重	車速	OFF
00006	A	コンピュータ	11:00:43.32	23.5	CD/MD	3	外気	INFO	車速	OFF
00006	J	INFO	11:00:47.87	23.5	CD/MD	3	外気	INFO	始動	OFF
00007		スタートキー	11:00:47.87	23.5	CD/MD	3	外気	INFO	始動	OFF
00007	L	フォロント デフ	11:01:30.37	23.5	CD/MD	3	外気	INFO	始動	ON
00008	B	リア デフ	11:01:30.88	23.5	CD/MD	3	外気	INFO	始動	ON
00008		スタートキー	11:01:30.88	23.5	CD/MD	3	外気	INFO	始動	ON
00008	X	風番	11:01:45.87	23.5	CD/MD	3	外気	INFO	始動	ON

図3 解析処理画面
Fig3. Analysis page

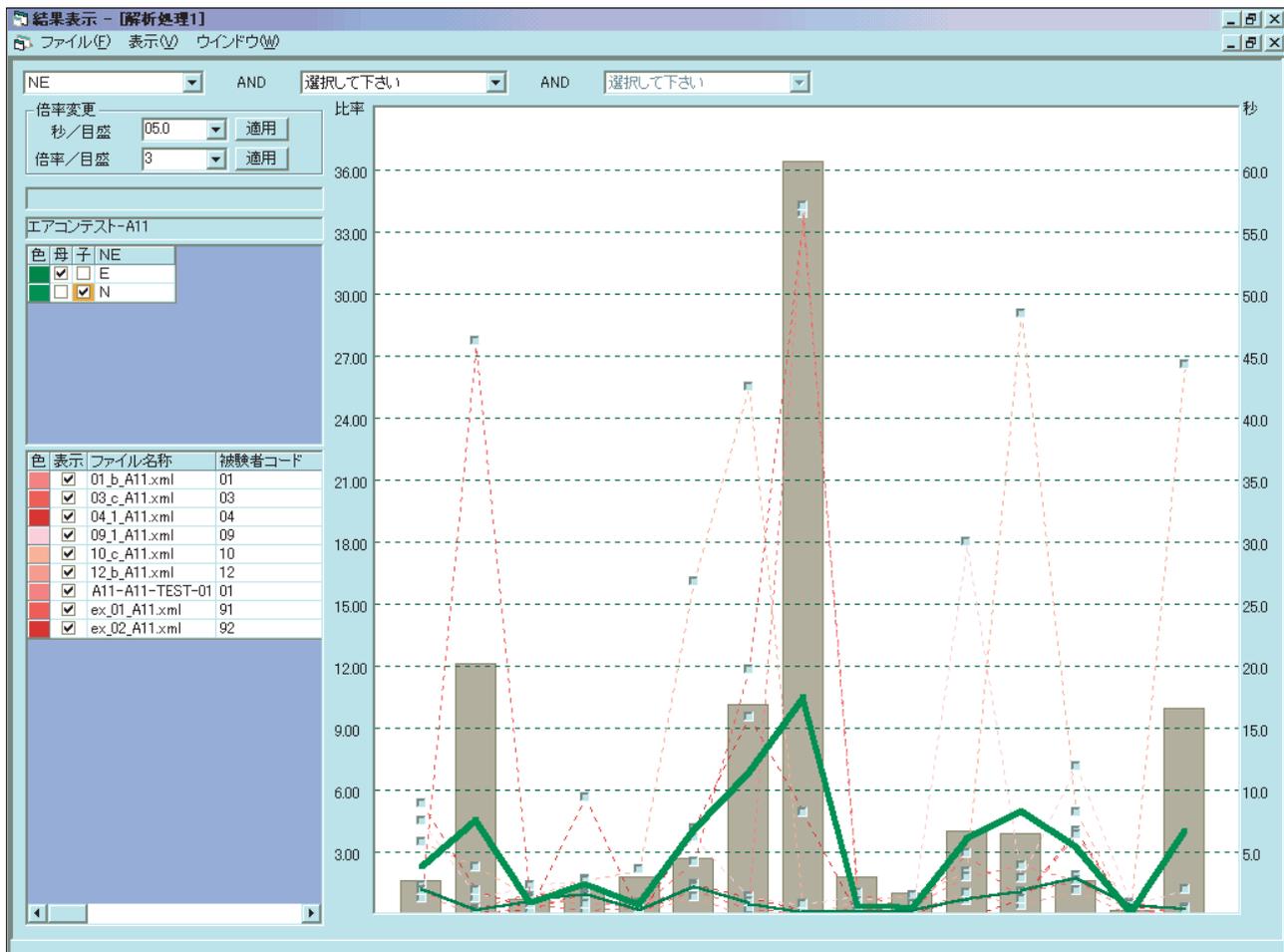


図4 グラフ表示
Fig4. Visualization

ることが可能になる。

本評価ツールはもともとNE比のみを比較するツールとして計画されたが、前述の考察の基に被験者情報で登録されたユーザ属性もNE比のような要領で比較できるように改良を行った。ユーザテストによって、同じターゲットユーザでも被験者として有効な場合もあればそうでない場合もある。これは現在のユーザ属性の抽出や整理、分析が不足していると考えることができる。

自動車の運転でいえば、運転経験年数や運転頻度、搭載機器の違いなど被験者の操作に対するモチベーションを測る属性は多岐に渡るであろう。

また、携帯電話でも経験した操作デバイスの種類や電子メールの活用頻度、入力文字数や定型文の活用度合いなどが見逃せない要素であろう。

こうしたユーザ属性毎に操作時間の比較を行えば様々なギャップが外化できると考えている。

7. 今後について

操作時間を計測してパフォーマンス評価や学習率を計測する手法は広く知られるようになったが、NEMのように問題発見に活用しようとする試みは少ない。本ツールは開発されたばかりのため現時点で報告できるのは、①業務効率の向上について、②考察力の向上、の2点が

中心となるが、本ツールの活用によってユーザ属性毎の様々なモデルギャップの存在が予測されており、この成果については次の機会に報告を行いたい。

今後の展開として、専門家でなくても分析可能なツールが実現すれば、開発セクションや品質評価部門での操作性検証がより容易になると考えられる。現状の運用体制化においても企業の品質評価指標として操作性が確実に検証できるようになれば、開発プロセスに操作性の検討工程が組み込まれることが期待できる。また、こうした指標がいずれ流通にも普及し店頭における操作性表示が実現すれば、ユーザが使いやすさで商品を選ぶ日もやって来るであろうと考えている。

8. 参考文献

- [1] Urokohara,H., Tanaka,K., Furuta,K., Kurosu,M., : NEM: Novice Expert ratio Method A usability evaluation method to generate a new performance measure;CHI 2000 Interactive Posters(2000)
- [2] 鱗原、田中、古田、黒須：設計者と初心者ユーザーの操作時間比較によるユーザビリティ評価手法；ヒューマンインタフェースシンポジウム'99 (1999)