

眼球運動とマウス軌跡によるWEBアンケート回答時の 確信度評価とその応用

人間環境学専攻 2010 年 3 月修了 47086692 緒方 明子

指導教員 岡本 孝司 教授

Thinking processes and interest may be related to a track of eye. Moreover, the track of eye can be used also for detecting the indecision behind a final reply to a query. Therefore, taking decision making on WEB survey as an example, the decision making mechanism under the network environment was investigated by observing the track of eye. A correlation between a track of eye and the track of mouse was also estimated to investigate a relation between a track of mouse and a mentality on decision making. In experiments, a question was answered in a multiple answered query and by a slide bar. The eye and mouse track in the multiple-query can be used to evaluate the value of the bar as a confidence factor.

Key words: web survey, track of mouse, eye movement, confidence factor, thinking process

1 序論

顧客の満足度やニーズ, 行動履歴等を把握することは企業にとって重要であるため, 既存手法に比べ迅速かつ低コストで大量のデータを取得可能な Web アンケートに注目が集まっている. 一方 Web アンケートによって詳細な意識調査を行うためには複雑かつ複数の設問が必要であるがこれらは回答者に負担がかかってしまいデータ回収率を低下させてしまう. 逆にシンプルな質問方式・質問数を採用すれば詳細な情報把握が難しくなってしまう. このように, 回答負荷とデータ回収率との間にはトレードオフの関係があり, 詳細な意識調査は容易ではない. 王子¹⁾は複数選択型問題(以下 MA)回答時のマウスの軌跡を利用して, 同時に行った順位付け型問題の結果を推定する方法を提案したが, 推定には経験的な指数関数や閾値が用いられ, またマウスの動きがそもそも何を表しているのかが明確ではなかった. そこで本研究では, 視線計測器を用いて回答プロセス中の視線の動きを測定・分析すると共にマウスの動きとの関係性を明らかにする. また, 詳細な意識調査を可能にする回答者毎の意見の強さを表した確信度の評価に利用する提案分析モデルの閾値を持つ意味を確認し, その最適化を行う. 更に離散値である順位よりも詳細に確信度を表すため, 回答者ごとに任意の値を取り得るスライドバー形式の設問を用いて確信度を調査し, その結果の推定を試みる. 以上により, 本研究ではシンプルな設問形式である MA に対する回答時の挙動から詳細な情報, すなわち確信度を評価する手法を開発すると共にそれら手法の応用性を Web アンケート評価実験により検討する.

Q1. リップに限らず, 普段どこで化粧品(全額)を購入しますか? (複数回答可)	Yes ○	No ○	上司にリーダーシップがあると感じる. 上司は良いリーダーであると感じる
<input type="checkbox"/> マウスオフの状態(透明度100%)	Yes ○	No ○	浜田選手が充電場所は風呂の方, 地域住民から好評であると思う
<input type="checkbox"/> マウスオンした状態(透明度0%)	Yes ○	No ○	チームが一丸となって業務に取り組んでいると感じる
<input type="checkbox"/>	Yes ○	No ○	自分の業務に使命感, やりがい, 喜びを感じる
<input type="checkbox"/>	Yes ○	No ○	自分の努力, 誠意や成果を正当に評価されていると感じる

Fig.1 Display design of Web questionnaire experiment1 (left) and Web questionnaire experiment2 (right)

2 Webアンケート評価実験 I

2.1 実験概要

実験では 19~23 歳の女子大生 20 名を対象として趣味・思考を問う内容で構成された化粧品に関する Web アンケートを実施した. Web アンケート回答時の視線の動きを TOBII 製 Tobii X50 で, マウスの動きを Ajax により測定し, MA に答えてもらった後, 同様の設問内容で各選択肢に対する意識の度合を 0%~100% で問うスライドバー式の設問(以下 SL)に回答してもらうという流れで MA と SL を 1 組とする計 33 ページの Web アンケートを行った. MA の画面デザインを Fig.1(left)に示した. 本研究では, SL の回答を正解(=回答者の確信度)とみなして, MA 回答時の挙動からこの正解を推定する. 以下, SL の値を「正解値」と呼ぶ.

2.2 実験データ

Web アンケート評価実験 I ではマウスの位置情報と回答情報を JavaScript によって取得し, Ajax よりサーバへデータを送信, perl を用いてログデータをデータベース化するアンケートシステムを構築している. それらで取得したマウスデータと視線計測器のデータを利用したある被験者のマウス(赤点)と視線(青点)の軌跡のトラック結果を Fig.2 に示す. また, Fig.3 はマウスと視線の y 方向座標時系列データのグラフを表している. これらの結果から, 1. マウスは概ね視線に追従して動く, 2. 問題文を読む・選択肢を一望する際には視線のみが動く, 3. 意味のないマウスの動きや欄外への停滞は無く, マウスは選択肢付近を移動するという 3 つの特徴が見られた. 設問文の理解に要した最初の部分以外, 視線が遷移した後にマウスがそれを追いかける形になっていた. 視線とマウスには遷移のタイミングのずれが見られるものの, 二つの挙動は良く一致していることが明らかとなった. また, マウスと視線間の距離は回答開始から終了までの全回答時間内では, $19.70(\pm 4.75)$ pixel, 最初にチェックを入れたときから最後にチェックを入れるまでの回答時間内では, $8.44(\pm 7.25)$ pixel であった. これは 30 pixel の範囲内に 100ms 以上滞在するという注視の定義から, マウスと視線間距離が注視範囲内であることが分かる. なお本研究における選択肢幅は 80 pixel であり, これはマウスと視線のずれや 30 pixel よりも大きくなっている. これらの結果から, 回答中の視線とマウスの軌跡は類似しており, マウス

の軌跡は視線の動きをある程度模擬していると考えられる。

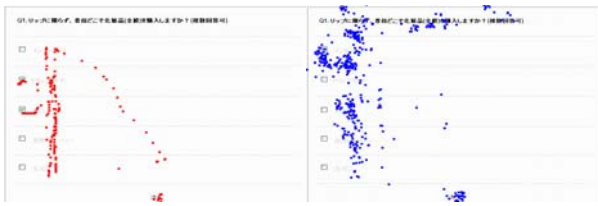


Fig.2 Trajectory of mouse(red) and eye movement(blue)

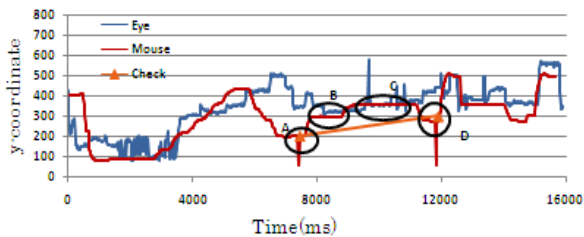


Fig.3 Time series graph of mouse and eye movement. A and D are expressive of checking option-2 and option-3. B and C show staying between option-3 and option-4.

2.3 分析モデル

従来のアンケート形式では 1(Yes)か 0(No)の二値データによって意思表示されるが、実際には回答者の意思は 1 から 0 の範囲のある数値(潜在得点)をもち、回答者独自の判断基準(閾値)を境に選択するかしないかが決定される。本研究では、迷いを表すとされる視線とマウスの動きが一致していたため、マウスの座標はどの程度その問題を検討しているかを表す指標であると仮定出来る。そこで、視線とマウスの軌跡の密集の度合いから 6 つのモデルによって潜在得点を算出し、各被験者の回答確信度を評価した。今回提案した分析モデルは“マウス加速度モデル(Ma)”, “視線加速度モデル(Ea)”, “マウス滞在時間モデル(Ms)”, “視線滞在時間モデル(Es)”, “注視時間モデル(Et)”, “注視回数モデル(En)”の 6 つであり、加速度モデルはマウスと視線の加速度が一定値(A)以下の条件で、ある選択肢付近に滞在した時間、滞在時間・注視時間はある選択肢付近に滞在・注視している時間、注視回数はある選択肢付近を注視している回数を利用して潜在得点を算出する。今回、加速度に着目した理由として、Fig.3 の結果から回答直後・終了時・選択肢間を移動するときには通過する選択肢について考えているわけではないということ、また特定のターゲットへ向かって移動するときには加速度が大きくなるということから、加速度を利用して無駄なマウスと視線の滞在時間を除去すれば、より詳細な分析が可能だと考えたためである。そのような時間の合計である条件付経過時間(t_n)が長い場合、回答者は判断に迷ったことになり、その選択肢に対する潜在得点は判断基準である閾値に近い値をとることになる。逆に、 t_n が短い場合、被験者は迷うことなく即決したことになりその選択肢に対する潜在得点は二値データとして 0 または 1 に近づくことになる。上記の仮定を満たすように、回答者が n 番目の選択肢に対して持つ潜在得点 T_n を決定する関数 $F(t_n)$ を(1)のように

定める。上式は結果的に選択した場合、下式結果的に選択しなかった場合の潜在得点関数である。また a, b, c は下記(*)を満たす定数で、判断基準となる閾値は 0.5 で一定とする。

$$T_n = F(t_n) = \begin{cases} \exp(at_n + b) + c \\ -\exp(at_n + b) + c \end{cases} \quad (1)$$

$$* \exp(at) = \exp(b) = c = 0.5$$

また今回、視線の加速度の閾値(A)を変化させて算出した潜在得点のずれ値(潜在得点と正解値の差)を分析し、視線と加速度の閾値(A)の最適化を行った。全ての選択肢上での加速度と結果的に選択された選択肢上の加速度の 4 種類で感度を分析した結果、閾値は全体としてずれ量の小さい 20%(全体の加速度から早い 20%を切った値)を利用することとした。

2.4 評価実験 I における分析結果

総合的に 6 つのモデルの有用性を検証するため、パー値と潜在得点とのずれ(差)の大きさを比較した。ずれが小さいほど分析モデルの当てはまりが良いことを示し、詳細な意識調査が実現可能なモデルとして期待出来る。Fig.6 は各問題ごとの全被験者の潜在得点と正解値の平均ずれ値を表している。Fig.4 からマウスと視線の加速度モデルは他のモデルと比べてずれ値が小さく、約 5%前後のずれ値で確信度を評価可能であるとの結果が得られた。また Table1 はモデル毎のずれ値平均とそれらの SD を示しており、マウス・視線の加速度モデルにおけるずれ値平均・SD はその他の分析モデルより当てはまりが良いことを確認出来る。

	Ma	Ea	Ms	Es	En	Et
ずれ平均	1.1	1.9	25.9	26.2	32.7	30.9
SD	0.8	1.2	7.7	9.2	7.9	7.9

Table1. average of deviation and standard deviation

これらの結果から、提案した 6 つの分析モデルにおいて最も確信度評価に効果的なモデルは加速度モデルであるといえ、マウスの軌跡と視線の軌跡は等価であるという結果から、マウスの加速度モデルのみでも同様の結果が期待出来る。実用的でない視線計測とは異なり大規模なネット調査であっても本手法を適用でき、Web アンケート評価実験 II では大規模アンケートに本手法を適用した事例を示していく。

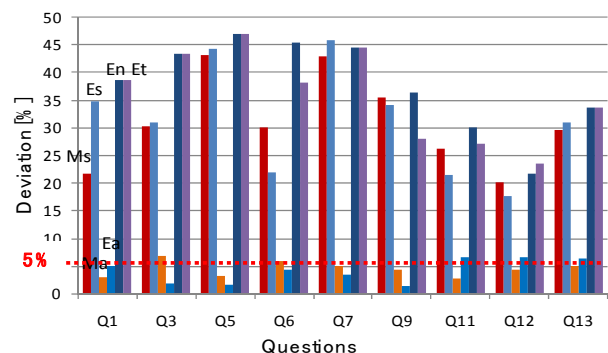


Fig.4 Deviation of implicit score and bar score

3 Webアンケート評価実験Ⅱ

3.1 目的

Web アンケート評価実験Ⅰの結果から、6つの分析モデルのうち、マウスと視線の加速度モデルによって分析された潜在得点が確信度評価に有効的であるということを示唆する結果が得られた。Web アンケート評価実験Ⅱでは、対象者数530人(有効回答者数:519人)からなる大規模Web アンケート実験を行い、ページデザインや質問内容の異なる環境下でマウスの加速度モデルの有用性を検証することを目的としている。更に、効果的なページデザインの検討や詳細な分析を行うための回答行動に対する工夫、確信度評価を行うための新たな分析手法の検討なども行い、詳細な意識調査を可能にするWeb アンケート手法の開発とその応用性の検討を行う。

3.2 実験概要

中部電力浜岡原子力発電所に勤務する一般職・管理職の社員計530名を対象とし、安全文化醸成に関するWeb アンケートを実施した。安全文化醸成活動の状況を把握するものや安全文化の4つの要素であるコンプライアンス、技術力、コミュニケーション、士気・やる気に関連する質問内容で構成されたアンケートを実施した。評価実験Ⅰと同様にAjaxを利用してマウスの回答情報を記録している。回答行動に対する工夫として、配置された選択肢の順序によって回答に偏りが現れないよう被験者毎に選択肢をランダム配置している。ページデザインは従来の選択型形式であるノーマルデザイン(MA)、Fig.1の右図にあるYesとNoボタンが隣に配置されているデザイン(YN1)、YesとNoボタンが選択肢を挟んで配置されているデザイン(YN2)の3種類で検討し、従来型MAで選ばれなかった非選択項目に関してもNoとチェックを入れてもらうことで非選択項目に対しても確信度評価を行う。ページ構成は評価実験Ⅰと同様、各ページデザインとそれらと同様の質問によって選択肢に対する意見の程度や割合を0~100%で聞くSLの設定を1組とする計23ページのWebアンケートとなっている。

3.3 評価実験Ⅱにおける分析結果

評価実験Ⅱでは、Yes/No型のデザインページを利用しているが回答挙動等の結果から、この形式では各選択肢を一つの設問とみなし回答している様な傾向が見られた。つまり、従来のように、選択肢間の比較を加味し回答を行うのではなく、その選択肢単体でYesかNoかを選択している可能性がある。そのため、新たな分析手法としてある設問におけるチェック間の時間間隔から確信度を評価する手法等も検討しているが、本要旨では潜在得点に関する分析に重点をおき説明していく。今回は、3種類のデザインページ毎にマウスの加速度モデルを適用し分析を行った。まず、回答者数73名のMAにおける分析結果をFig.5(右図)に示す。各ページの選択肢毎の平均潜在得点と平均正

解値の分布を表しており、ほぼ45度線上に分布していることが分かる。これら全ページにおけるずれ値平均は3.7、SDは1.6となっている。これらの結果は回答論理矛盾(選択した設問の正解値<選択肢しなかった設問の正解値)の回答者を不正回答者として除いているが、チェックした項目だけでなくNoを選択した回答者全て含まれた結果である。

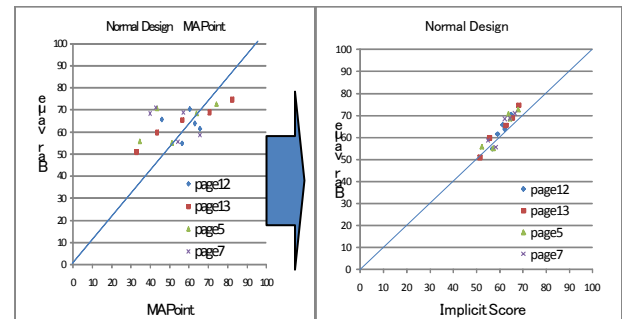


Fig.5 Bar value and MA point (light) or Implicit score (left) of Normal version

また従来の集計法である単純なMA集計ポイントで各ページの選択肢ごとに算出したMAポイント(ある選択肢における累計チェック数/そのページの回答者数)と正解値との関係性はFig.5(左図)となっており、平均ずれ値は11.4、SDは9.5であった。ずれ値の算出式は $dn=(Jn-Tn)/Tn$ となっている。Jnはある項目における潜在得点、Tnは正解値を表している。また、潜在得点とMAポイントの各ページの全選択肢における標準偏差(SD)をTable2に示す。

	page12	page13	page5	page7
MA Point	7.7	19.9	15.9	10.6
Implicit Score	3.0	7.0	6.3	4.6

Table2 standard deviation results of MA Point and Implicit Score

Fig.5、Table2の結果から、従来型の集計法であるMAポイント値の分布より本分析モデルを適用した結果、乖離が小さくなっていることが分かる。同時に潜在得点の当てはまりの良さを確認することが出来る。次にYN1、YN2における分析結果をFig.6に示す。

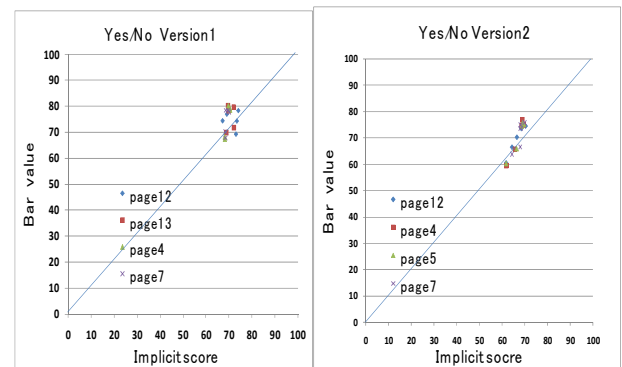


Fig.6 Bar score and Implicit score of Yes/No Version1 (left) and Yes/No Version2 (right)

これらの結果は全てのページ及び選択肢の平均値が潜在得点・正解値ともに 60 以上となっている。実際にこれら 4 つのページでの Yes 選択者の平均比率はデザインページ 1 が 84.7%, デザインページ 2 では 72.5% となっており, 大多数の被験者がほぼ全ての選択肢にチェックを入れていることが要因となっている。この様な分布形態は質問内容に依存するものの, 今までと違い No という明確な意見を問う質問形式であったため回答者にとって No に対する敷居が上がり過ぎたことが原因として考えられる。また, 分布形状に着目すると, Fig.6(left)は一箇所に各選択肢が集まっており一見正解値との当てはまりが良好でない様に見られるものの, 正解値との平均ずれ値は 5.5, SD は 3.9 となっており, モデルの適用効果自体に大きな影響は見られなかった。Fig.6(right)の平均ずれ値は 1.35, SD は 2.5 となっており, これら二つのデザインページにおいても加速度分析モデルが効果的に適用されていることが分かる。また, これらのページの MA ポイント値と潜在得点両者における正解値とのずれ値を比較した結果を Table3 に示す。A, B, C は MA, YN1, YN2 の潜在得点におけるずれ値, MA_A, MA_B, MA_C は MA, YN1, YN2 の MA ポイント値におけるずれ値を表している。これらの結果から, 2 カ所の選択肢でモデル適用後にわずかにずれが大きくなっているが, 殆どの場合, 加速度モデルに当てはめた潜在得点の方がずれ量が小さいことが分かる。また, Web 評価実験 I では選択された項目のみずれ量を検討したが, Web 評価実験 II では非選択項目も加味している。選択されていない項目, つまりスライドバーの値が高くなく順位が下位のものであってもある程度の精度で確信度評価が行えることが分かった。

	item1	item2	item3	item4	item5	average
A	54	21	43	15	24	31
B	73	79	38	42	09	48
C	51	22	12	42	38	33
AMA	98	45	195	06	19	73
BMA	118	142	157	179	157	151
CMA	159	37	68	144	162	114

Table3. Deviation and Average of MA point and Implicit Score

次に推定モデルがどの程度少ないサンプルサイズまで効果的に機能することが出来るのかを検証した。Fig.7 の Implicit-Score2(青線)と MA-Point2(赤線)はあるページにおける潜在得点と MA ポイント値のずれ値をサンプルサイズ毎に算出した結果を示している。MA ポイントは全体の平均であり, 60%前後程度のサンプルサイズで最終的なスライドバー値(正解値)の平均を予測することが出来る。サンプルサイズは 30%~100%(=73 人)を分析対象として無作為抽出を行った。サンプルサイズが 40%(29 人)前後までは比較的結果が安定していることから, 本研究のモデルはサンプルサイズが 40%程度以上から有効なもの

と考えられる。MA ポイント値と潜在得点のずれ値の差が開くのは 50%付近(36 人)からとなっており, 50%のサンプル数から潜在得点の効果を確認することが出来る。また Fig.7 の Implicit-Score(緑線)と MA-Point(紫線)はサンプルサイズを母集団ごと減らした際の潜在得点と MA ポイント値の当てはまりを確認するグラフとなっている。結果から, MA ポイント値はサンプル数を減らすほど, 当てはまりが悪くなるが, 加速度モデルで算出した潜在得点はサンプル数を減少させた場合でも継続的にモデルが機能することが分かった。

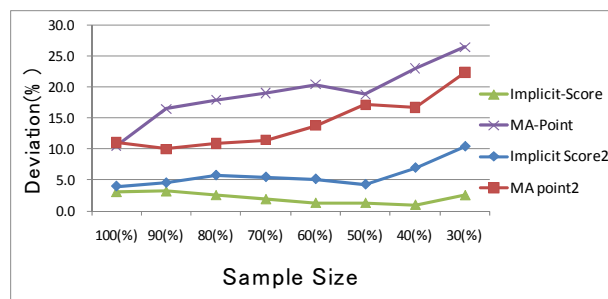


Fig.7 Causal relationship between Deviation and Sample Size

Webアンケート評価実験 II では加速度モデルの効果を検証するため被験者 530 人を対象に大規模Webアンケート実験を行い, これらの結果からデザインページや質問内容に多少の依存はあるものの加速度モデルが効果的に機能することが再確認出来た。また単純集計手法のMAポイント値と比較した場合でも本モデルの当てはまりの良さが分かり, 有用性を証明出来た。一方で, MAポイント値と潜在得点のずれ量が逆転する結果が一部得られたりと精度の向上に努める必要もある。その他スライドバーの回答挙動から選択/非選択の結果を分析や, 新たな分析モデル手法の提案なども課題として挙げられる。

4 結論

本研究では Web アンケート回答時における意思決定・認知プロセスから確信度を評価した。視線とマウスの動きの関係性を分析し, 両者の傾向が良く一致することを明らかにした。また, 視線とマウスの加速度による条件付き滞在時間から回答確信度を精度良く推定出来ることが分かった。更に質問内容やページデザインなど実験状況が異なる環境下で分析を行い, マウスの加速度モデルの当てはまりの良さを実証することが出来, 従来型の集計法と比べてもその効果が確認できた。これにより, マウスの加速度によって回答者の迷いを抽出した詳細な意識調査が可能となり, 回収率と情報量・信頼性などのトレードオフ問題を解決できる Web アンケート手法の提案を行えたといえる。

文献

- 1) 王子富幸 : Web 環境下における意思決定に関する研究, 東京大学修士論文(2007)