

## 確信度理論の展望

### A Review of Confidence-Rating Process Theories

妻 藤 真 彦

本稿における確信度 (confidence value) とは、心理学実験において被験者が行った反応 (記憶の想起や質問への解答など) の正確さを、被験者自身が自己評価した結果のことである。殆どの実験において確信度は数値で評定される (例えば1点から7点までの数値で評定する7段階尺度など)。

これまで確信度は、様々な認知に関する理論を検証するための重要な測定値として使用されてきた。また「自己評価能力」の研究では、確信度そのものが研究テーマである。そして、それら個別の研究の背後に仮定されてきた理論がある。この理論では、人間が認知・想起・思考などの情報処理を行うとき、そのときの認知処理システム出力の「強さ」(strength) がモニターされており、モニター結果を評定値として表現したものが確信度だと仮定している (記憶などでは、「強い出力」が記憶システムから出てきたときは、その内容について、より高い確信度を答えることになる)。本稿で詳しく述べるように、この仮定は極めて重要なものであり、そしてまた、大きな問題を含んでいる。本稿の主目的は、これが基本仮定としてどのように受け入れられてきたかを概観し、その問題点とともに理論的展望を論ずることである。

#### 信号検出理論

再認記憶の研究では、確信度評定のメカニズムが信号検出理論の枠組みの中で理論化されてきた。ここでいう信号検出理論 (signal detection theory; e. g., Tanner & Swets, 1954; Swets, Tanner, & Birdsall, 1961) と

は、元々は感覚器官に与えられる刺激 (信号) の有無を検出する課題、つまり信号と何らかのノイズ (レーダーなどの機械又は脳自体から発生しているノイズ) を区別する課題での、反応確率を説明する理論である。このような実験では4種類の反応が区別される。つまり、正しく信号があることに気付く (hit)、あるいは見落とす (miss)、また信号はないのに、ノイズを信号と見誤る (false alarm) あるいは正しく「信号はなかった」と答える (correct reject) という4種類である。

強い信号なら見逃される確率は小さい。逆に弱い信号は見逃される率が大きい。ただし信号強度はいつも一定ではなく、偶然強くなったり弱くなったりする。またノイズは全般には信号よりも弱いですが、これも確率的に変動するため、最も強いノイズの方が、最も弱い信号よりも強くなってしまふこともある。つまり信号に似た強度になってしまう。そのため、弱い信号を受け取ったときには、それが信号なのかノイズなのかの判定は難しい。結局、受け取った刺激の「強さ」が、ある一定の値より大きければ信号であると判定し、それよりも小さければノイズだと判定する (つまり、信号はなかったと判定する) しかないのである。

Figure 1は、この関係を示す。判断系に入力された情報の「強さ」(strength) が図中の基準値 (criterion) よりも右であれば、yes (信号あり) と答え、左であればno (信号なし) と答える。実際に信号があるなら、信号強度分布のピーク付近の「強さ」が最も出現しやすい。ないときにはノイズ分布のピー

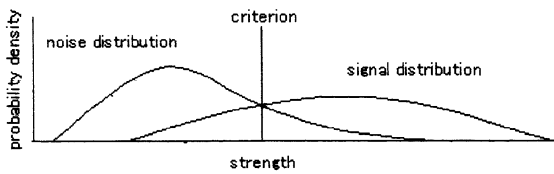


Figure 1 . The signal detection theory.

ク付近が出現しやすい。しかしノイズのみであっても、基準値より右の「強さ」になることもある。その確率（ノイズ分布の基準値より右の積分つまり面積）が false alarm の確率である。また信号分布の基準値より右の積分（面積）が、hit の確率になる。

この Figure 1 では、ノイズと信号各々の分布は異なる分散を持つように描かれている。この2つの分布の型が等しいとは限らないという意味で、数学的に一般的な形で描かれている。しかし具体的な理論として応用されるときには、しばしば、これらの2つは期待値の相違以外は等しい分布だと仮定されることも多い。また、この図では「強さ」が連続変数として描かれているため、縦軸が確率密度と表記されているが、離散確率変数であるようなモデルもある。その場合には、縦軸は単に確率である。

また、どのような強度を判定基準とするかは、各被験者によって、あるいは条件によって異なる。miss を少なくする方が有利であるような条件下では、false alarm を犯す確率が大きくなるのを甘受して、小さな強度値のところを基準値に選ばれるであろう。Figure 1 では、ノイズの確率密度と信号の確率密度の比（尤度比：likelihood ratio）が1になるところに基準値が置かれている。尤度比の値を用いて判定基準を表す場合、ベータ（ $\beta$ ）と表記される。

判定結果に影響を与える変数として、信号とノイズの強度の確率分布、特に重要なのは各分布の期待値（平均値）間の距離（ $d'$ ）である。これが大きいほど、当然判定は正確である。感覚刺激を検出するという実験の場合、これは刺激の選び方によって決まる変数である。もちろん、各分布の分散や型も miss 率や false alarm 率に影響する。

信号検出理論は感覚刺激の検出実験に当てはめられるだけでなく、様々な認知判断のメカニズムを説明する理論の枠組みとしても使用されてきた。その意味で、信号検出理論とは上記のようなメカニズムに関する一般理論（数学的性質に関する理論）である。個別の分野に応用されるときには、具体的なモデルとして、各種のものが提案されることになる。

例えば、いくつかの項目を記録するという実験課題について、よく覚えているということ、「記憶強度が大きい」（「強さ」が大きい）と表して、信号検出理論の特定のモデルを想定することになる。このとき、記憶強度を連続変数としてモデル化する場合、しばしば「信号」とノイズの分布両方とも正規分布あるいはガンマ分布だと仮定される。一方、各々の記録項目が持っている属性（その項目を他の項目から区別するのに必要な情報）が複数あるものとする理論では、記録された属性の数が信号検出理論の「強さ」に当たると仮定する。このような記憶理論での記憶の「強さ」は、例えば二項分布する不連続変数である。

このような理論は、確信度が何を測定したものであるかという、データ解釈にも応用されてきた。例えば、old-new 型（yes-no 型）の再認課題への適用が典型である。old-new 型では、単語や絵などの記録リストがまず被験者に呈示され、その後、単語（絵）がひとつずつ呈示されて、その都度、先に記憶した項目リストに含まれていたものであるのか、そうではないのかという判断が要求される。当然、記録リストの中に含まれていた項目であれば、「old (yes)」という判定が正答である (hit)。記録リストになかったものであれば、「new (no)」が正答になる (correct reject)。しばしば、解答時に確信度も答えるように要求される。

このような実験課題は、信号検出理論によれば次のように表現できる。すなわち、記録された項目が「信号」に相当し、再認検査時に初めて呈示される項目が「ノイズ」に相当する。そして、記憶過程からの出力が、なんらかの情報プーリング過程によって、まず一次元の量すなわち記憶の「強さ」に変換される (e. g.,

「熟知度」などであり、Figure 1 なら、横軸が記憶の「強さ」である。前述の記憶理論の例では、個々の記銘項目について、記銘されていた記憶属性の数が記憶の「強さ」だと仮定される。その「強さ」がある基準値 (Figure 1 の「強さ」を表す横軸上の点、*criterion*) よりも大きければ、被験者の反応は「old (yes)」になる。つまり被験者は (実験室で記憶するように指示された項目リストの中に、あるいは裁判での証言なら、事件の現場で) 「見たことがある」と答えることになる。当然、その「強さ」が基準値以下であれば、反応は「new (no)」, つまり「見たことはない」という答えになる。そのような反応に対してなされる確信度の評定も、解答の決定に使われるのと同じ出力量、つまり「強さ」(熟知度など) をモニターした結果だと仮定される。

この仮説の下では、ある被験者の old 反応に関する確信度が大きかったなら、その評定が基づいていた記憶の「強さ」は大きかったのだと解釈される (厳密に言えば、そのときの想起の強さは基準値よりもかなり大きかったことになる)。ただし、new 反応 (「見たことはない」) の場合には、上記とは逆に、記憶強度が基準値よりも小さい (強度と基準値の差が負の値である) ときに生ずる反応であるから、強度がより小さいほど、それをモニターして生ずる確信度は大きくなる。とはいえ、基準値から離れているほど確信が強いという点では old 反応の場合と同じことである。そこで、記憶強度と基準値の差の絶対値をモニターした結果を、数値で答えたものが確信度だという言い方もできる。

言い換えると、ある「old (yes)」反応が高い確信度を伴っているならば、そのような反応は、次のような実験条件での「old (yes)」反応と同等だったのである。

その実験では、false alarm (間違った「old (yes)」反応) を減らすために、厳しい基準つまり記憶強度の大きな値に基準値をセットするように被験者は誘導されている。要するに、Figure 1 の基準値よりもかなり右の方に移動させるのである。誘導の方法として、例

えば、各試行での反応の正誤に対して被験者に報酬が与えられたり、あるいは損失を被るように実験手続きが設定されている。そして、正答の「old (yes)」反応 (hit) に対する報酬と比べて、誤答の「old (yes)」反応をしたときの損失がより大きいように配分されている (hit なら一回につき5円もらえるが、false alarm なら10円とられるなど)。このような場合、誤った「old (yes)」反応を減らすように、被験者は基準値を、より「強い想起」の方に移動させるであろう (つまり、かなり強い熟知度でない限り、old とは答えないことに決める)。もし、確信度が想起の「強さ」(熟知度) をモニターした結果であるのなら、このような誘導条件での「old (yes)」反応が、そのような基準点誘導のない実験での、大きな確信度を伴う「old (yes)」反応と同等だということになる。

つまり、大きな (小さな) 確信度を伴う old 反応は、大きな (小さな) 基準値をセットするよう誘導されたときの old 反応と同じことである。この仮説の証拠を得るために、基準値誘導のない実験について、hit 率や false alarm 率が各確信度の値ごとに集計され、それらが、基準値誘導実験のデータと比較されてきた。要するに、そのようなデータは、幾つかの異なる基準値 (を誘導した実験条件) ごとに集計された hit 率や false alarm 率と同じ傾向を示すはずである (これらをプロットしたグラフは ROC 曲線, receiver operating characteristic curve, と呼ばれている)。この予測は繰り返し確認されてきた (e. g., Bernback, 1967; Donaldson & Murdock, 1968; Kintsch, 1967; Murdock & Dufty, 1972; Norman & Wickelgren, 1965; Murdock, 1974, for review)。そして、このような信号検出理論のパラダイムに基づいて、より詳細な確信度評定の理論が提案されてきた (Wallsten & Gonzalez-Vallejo, 1994)。もっとも、Humphreys, Bain, and Pike (1986) は、記憶の再生については all or nothing 的であり、再認出力のみが何らかのタイプの量的なものだと主張してはいる。

## 確信度評定の標準仮定

従来、確信度に関して、被験者の反応に関する被験者自身の確信度が大きかったなら、基本的にその反応が基づいていた記憶の「強さ」が大きかったのだと仮定されてきた（再認などのように yes-no 型の課題では、その「強さ」と基準値の差の絶対値が大きい）。つまり、前節で詳述した、確信度を信号検出理論と関係づける仮定である。知覚や記憶の数理モデルを検証する際に、この基本仮定が当然の前提として組み入れられたり、あるいは実験データを分析するときの、分析手法として使われることも少なくなかった。このような事情を以下に述べるような例で示そう。

Murdock (1974) は (p.26-28)、この確信度に関する基本仮定を、記憶理論の検証方法として利用できると述べている。すなわち、被験者が反応するたびに、その反応に対する被験者の確信度をとっておけば、単に正答か誤答かという測定値だけではなく、そのときの記憶出力の「強さ」も測定できたことになるというのである（もっとも「そのように仮定して良い」という言い方ではあるけれども）。Murdock はこれを「Model Tests」という節で述べているが、そこで言う Model Tests とは、この確信度に関する「仮定」のテストではない。そうではなくて、記憶モデルのどれが、より優れているかをテストする「実験方法」を論じた部分なのである。言い換えると、「仮定」という言い方は為されていても、それは各種の記憶モデルからデータを予測するときに、各々のモデルすべてに、この確信度に関する「仮定」を組み込み、そうして得られた予測を検証するという、「研究法」として述べられているのである。さらに、別の章では (p.119)、確信度を用いることで、「記憶の強さと判断過程（基準値の設定）を分離した考察を導くことができる」とすら書かれている。

Barakrishnam and Ratcliff (1996) は、視覚図形を分類する課題や再認記憶課題に共通の判断メカニズムについて、対立する2つの理論をそれぞれ表現する数理モデルを作り、さらに両方のモデルに、確信度に関する上述の基本仮定を組み込んで確信度と反応率の関係

を導き、そうして得られた予測値をデータと比較している。

また、Glanzer, Adams, Iverson, and Kim (1993) は、記憶に関する数理モデルを検証するにあたって正答率と確信度を用いているが、確信度については特別な断りもなしに、確信度を予測する式をいきなり書いている。その式を解説すると、上述の基本仮定が前提になっていることがわかる。Murdock (1974) は、確信度の値が異なる記憶強度（信号検出理論での「強さ」）に対応していると断っているが、ここでは、もはやそのような断り書きもないので、90年代には、基本仮定が既に当然の常識として通用していることが分かる。

再認記憶のメジャーとして、須藤 (1993) は確信度を用い、平均確信度の差をもって、記憶表象ネットワークモデルからの予測を裏付けている。この場合は、明らかに、表象ネットワークの「活性化」の「程度」（「強さ」）が「確信度」の大きさで「表現」されていると「仮定」されている。しかし、その「仮定」についての断りも、文献を挙げてのディフェンスも行われていない。Maddox and Estes (1997) は、かれらの記憶モデルの証拠として、須藤 (1993) と同様に、確信度の平均値をいきなり用いている。さらに、いくつか文献を挙げた上で、異なる確信度は各々異なる基準値に対応しているという前提を置いたときに可能になるもう一つの統計手法を用いた結果も示して、自説の根拠としている。

もちろん、ここでいう基本仮定だけで、すべての確信に関するデータが説明できるというのではない。しかし、最も基本的なところに、この仮定が用いられているのである。例えば、Griffin and Tversky (1992) のように、確信度が、記憶や、推論における証拠の「強さ」に依存するのは「当然」として、ただし、もう一つの要因として、その情報のソースに関する「重みづけ」あるいは「信頼度」にも影響されて、最終的な確信度が決まるという理論もある。このように、個々の研究分野では、様々な仮説が提案されてはいるが、この Griffin らの論文でもそうであるように、確信度の

基本仮定は、やはり「基本的」なものなのである。そして、前述のように、記憶などの分野では、しばしば当然のこととして、理論の検証やデータ解釈の前提に置かれている。

このような例から、確信度に関する基本仮定を「仮定」であると断らずに、確信度が（何らかの）「強さ」を表すものだという前提でデータを解釈するのは、かなり一般的だと言えよう。このような使い方は、例えば、ある電気回路の振る舞いがどうなるのかを測定によって検証するときに、電圧が電流と抵抗の積に等しいという「法則」に基づいて「目盛り」が付けられた測定器を使うことに等しい。つまり、この確信度に関する仮定は、「仮定」あるいは「仮説」というよりも、むしろ確立された「法則」として扱われているのである。

#### 標準仮定の問題点

しかしながら、Saito (1998 a) は、この仮定に関する実験的検討が、まだ十分に行われてはいなかったと主張し、この仮定の妥当性を検証するために新しい実験法を考案した。この論文では、検討対象を「再認や推論過程の出力が次元の量に変換され、その量を直接モニターした結果が確信度である」と表現し、これを標準仮定と呼んだ。ただし、この量は、例えば old-new 型の課題では、ある再認反応のもとになった情報の「強さ」と判断基準値の距離である (Figure 1 参照)。ここで言う情報の「強さ」は、個別の研究分野では様々な名称で呼ばれている。例えば、記憶系の出力だけが関与しているような実験の場合には、「幾つも出力されてきた記憶属性についてプーリングが行われる (もっとも単純なモデルでは、出力できた記憶属性の数が数えられる)」という理論になるが、これに例えば「熟知度」あるいは「再認強度」という用語が当てはめられるであろう。また、記憶系の出力だけではなく、様々な知識に基づく推論も同時に行われているような場合には、記憶系出力と推論系出力を全てプーリングして得られる量 (最も単純なモデルであれば、出力された記憶属性の数と、推論の元になった根

拠の数あるいは根拠の「強さ」の合計) が判断に使われるという理論になり、その量は「根拠の強さ」などと呼ばれるかもしれない (strength of the available evidence, Griffin & Tversky, 1992)。

しかし、本稿では、このような記憶の「強さ」や根拠の「強さ」などを、単に「強さ」と表現する。なぜなら、ここでは、特定の記憶や推論を説明する理論の検討が目的ではないからである。そうではなくて、様々な理論に共通に使われている、確信度評定過程に関する標準仮定が検討対象なのである。すなわち、例えば old-new (yes-no) 型の再認課題では、この「強さ」がどのくらい基準値より大きいか (old 反応の場合)、あるいは小さいか (new 反応の場合)、言換えれば、その「強さ」と基準値からの距離 (差の絶対値) がモニターされ、その結果を、数値を用いて表したものが確信度評定値だという仮定である。Saito (1998 a) の実験では、2 肢選択式の一般知識問題を被験者の課題として用いたけれども、この知識問題に解答するときの記憶系・知識系のメカニズムについては、特定の理論を前提にしたり、あるいは特定の理論的立場を執ったわけではない。したがって、ここでは記憶系、知識系、推論系などの「出力」がプーリングされたものを、特定の理論に依存する名称で呼ぶことを避け、単に「強さ」などと表現する。

標準仮定は、具体的に確信度評定に関わる情報処理過程を記述する理論ではなく、そのような理論の前提になるものである。そのため、個々の具体的な理論の中では、上記の (記憶システムや推論システムからの出力の) 「強さ」が、「熟知度」や「根拠感の強さ」などと呼ばれたりするだけではなく、この標準仮定自体も、各々異なる表現がなされている。

また各確信度の差 (各評定尺度間の距離) について、明確な仮説はない。つまり、例えば 7 点尺度が使われたとしても、ある実験における 4 点と 5 点の違いが、別の実験での 4 点と 5 点の違いと同じ「内的確信」の相違を示しているのかどうかについて、明確な合意はない。Stretch and Wixted (1999) は、 $d'$  (Figure 1 での、2 つの分布の距離) が異なる 2 条件を比較し

たとき、 $d'$ が小さい方の条件で各確信度値間の距離にファンアウトが起こると解釈できる結果を得た。ここでのファンアウトとは、 $d'$ が大きい条件のときの（例えば）確信度4と5の幅（心理的差）よりも、 $d'$ が小さいときの4と5の幅（心理的差）が広いということである。もっとも、数理モデルによる数値的予測はうまくいっていない。ただし、これは標準仮定が正しいとしたときのデータ解釈であり、また各条件内のヘテロジニアティはゼロだと仮定されている（ヘテロジニアティについては、妻藤，1999 aを参照されたい）。

この問題は、old-new (yes-no) 型の課題における基準値の位置 ( $\beta$ ) に関する理論上の議論とも深く関わっており、特にペイオフ上の偏りが無い課題では、基準値は尤度比が1になるところに、常にセットされるという尤度比モデル (e. g., Glanzer, Adams, Iverson, & Kim 1993; このモデルでは Figure 1の横軸は、元の分布から計算される尤度比である) や、ある1点に、ほぼ固定されているというヒューリスティックモデル (e. g., Barakrishnam & Ratcliff, 1996) の論争が続いている。

これらの論争点に対して、中立であるような標準仮定の検証が必要である。つまり、確信度値間の距離、および基準値の位置がどのようにして決定されているかという問題とは無関係に、標準仮定の本質のみを検証するための実験法が必要である。Saito (1998 a) の実験は、これらの問題を回避するように組み立てられている。

Saito (1998 a) では、集計において確信度を間隔尺度ではなく順序尺度として扱う統計法を中心的に使うことによって、確信度値間隔のファンアウト（ファンイン）現象が、もし起こっていても、それとは無関係な統計量で考察できるように工夫されていた。また以下で述べるように、基準値の位置問題については、2肢選択課題を使うことで回避している。

この実験において、被験者の課題は一セットの一般知識問題であり、2つの選択肢から一つを選ぶことで解答される48個の質問項目から成り立っていた。この

場合、標準仮定は次のように表現できる。まず、解答決定過程が2つの解答選択肢について、各々の「強さ (strength)」を比較する。この「強さ」は質問と選択肢について情報を処理する何らかの認知過程が出力する「情報」である（記憶系出力や推論系出力などのすべてがプールされた次元量を示す情報）。被験者によってなされる解答は、2つの選択肢から、より「強い」方が選択されることになる。以下では説明の便宜のために、2つの「強さ」の違いを表す量、つまり片方の選択肢の値から他方のものを引いた値を「知識差」と呼ぶ。つまり、信号検出理論における、2つの選択肢の「強さ」の比較は、2つの「強さ」の差をとることだと表現できる。ただしこのとき、その「知識差」の値の符号にはあまり意味がない。正であれば片方の選択肢が選ばれ、負であれば他方が選ばれるというだけであるため、この「知識差」の絶対値が確信度の理論にとっては重要である。標準仮定を、以下で説明するように、この実験法に沿って表現したものを、標準仮説と呼ぼう。

この実験において、特に2肢選択課題が用いられたのは、old-new (yes-no) 型の課題よりも、信号検出理論を簡明に表すことができ、また確信度の標準仮定もより単純に表現できるからである。そして、この方法によって、上記のような基準値の設定位置問題を回避することができる。Saito (1998 a) では、この点を詳述しなかったので、以下に詳細な根拠も述べよう。

もし、1つの選択肢について yes か no で答える課題なら、「強さ」がゼロに近い場合には、「基準値よりもはるかに小さい」、つまり基準値との距離が大きいため、「これではない」(no) という強い確信度になり、「強さ」が最大値に近いときには、これも基準値との距離が大きいため「これだ」(yes) という強い確信度になるという記述をすることになる。一方、2肢選択の場合、単にどちらの選択肢が強いかということだけ考えればよく、次に説明するように、基準値を考慮する必要はない。

一例として、一つの選択肢 A の「強さ」が非常にゼ

ロに近く（基準値よりはるかに小さく）、他方の選択肢Bもゼロに近いがAよりは強い（基準値よりは小さいがAほどではない）という場合を考えてみよう。このとき、もし基準値をいちいち考慮した記述をするなら、Aと比べてBの方が、「これではない」という程度が小さいためにBが選択されるという表現になる。また、Aは基準値より大きくBは基準値より小さいなら、Aは「これだ」と判定されるが、Bは「これではない」なので、Aが選ばれる。しかし、このようにAの「強さ」とBの「強さ」について、各々基準値との関係を考えても、実は両方の項に基準値が含まれるために、引き算すると基準値は消えてしまうのである。要するに、 $X$ を基準値とすると、「知識差」は

$$(A - X) - (B - X)$$

であるから、カッコ内が正の値（「これだ」という程度）であっても、負の値（「これではない」という程度）であっても、比較（差）においては単に $A - B$ になってしまう（ここでは選択肢Aの「強さ」を $A$ とし、選択肢Bの「強さ」を $B$ とした）。従って、いちいち基準値との関係を考慮した記述をしたとしても説明が複雑になるだけであり、「強さ」の差、つまり以下では「知識差」を使って議論を進める方がよい。2肢選択の場合には、基準値を考えた記述をする必要はないのである。

上述のように、old-new型の再認課題の場合、基準値の設定位置問題が、どうしても実験結果の解釈をあいまいにしてしまう。それに対して2肢選択課題であれば、実験結果の解釈にあたって基準値を考慮する必要がないため、確信度評定の本質的な問題点を直接検証することができる。

「知識差」を用いた表現法なら、次のように要点が整理できる。記憶システムは確率変動をともなっているので、「強さ」はある程度確率的に変動している。当然「知識差」もそれによって変動する。一つの解答が選択されたとき、「知識差」の絶対値がモニターされる。この絶対値を数値で表現したものが確信度である。このように、2つの選択肢に対応する「強さ」を個々に考えるのではなく、「知識差」を用いて、標準

仮定を2肢選択課題に特定の表現で表したものを、ここでは2肢選択課題の標準仮説と呼ぶ。この標準仮説は標準仮定の、「ひとつの表現」である。すなわちこの形式化は説明の便宜上のものであり、標準仮説に関する議論は、標準仮定を含むすべての具体的理論に当てはまる。

この判断時点での確率変動は、被験者の解答のゆらぎをもたすだけでなく、確信度自体の変動も引き起こし、そして、標準仮説が正しいなら、この確率変動によって特定の変動パターンが現れると予測された。

Saito (1998 a) は、同じ一般知識問題の実験セッションを2回繰り返して、セッション間の変動を検討した。その結果、解答の変動（ふらつき）率は確信度の単調減少関数になり、標準仮説に基づく予測と一致した。しかし、確信度の変動データは、深刻な問題を引き起こした。セッション間で同じ質問への解答が変化したときの、確信度の変動パターンが、予測されたものとは、あまりにもかけ離れた結果になったのである。

この最後の結果は、単に予測とデータの適合度が低いということではなく、セッション1で確信度最小であったときの、セッション2での確信度頻度分布のみが予測と一致し、それ以外、つまりセッション1で確信度が最小ではなかったデータのすべて（7つの確率分布中6つ）は、セッション2での確信度頻度分布が、予測とはかけ離れたものになった（セッション1での確信度が高いほど、予測から遠くなり、最大ときには予測と完全に正反対になった）。

この結果について、妻藤 (1998 b) は、被験者数を倍以上にして追試を行い、全く同様の結果を得た。ただし、Saito (1998 a) と妻藤 (1998 b) が使用した問題リストの内容は、すべての質問が同じ分野のものであった。つまり、ヨーロッパの都市名について、その所属する国名を選択肢とするものである。上記の現象が、このような地理に関する知識に特有のものではないことを示すために、妻藤 (1999 b) は、一般知識問題の内容のバリエーションを増やしたリストを用い

てさらに検討した。地理に加えて歴史、芸術、あるいは科学技術などに関する知識問題を混ぜたリストである。そして、ここでも同質の結果を得ている。もっとも、この問題リストでは、セッション間の解答変更率は、上述の2つの実験よりも少なく、また各確信度に対する正答率の関数、および各確信度に対する解答変更率の関数も、全体に小さい値へのシフトが見られた。

このように、新たに発見された現象は、再現性および記憶内容に関する一般性を持っていると判断されよう。Saito (1998 a) は、標準仮説に代わる理論として、「ふらつき」仮説（あるいは「迷い」仮説）を提案した。この仮説では、再認の強度に応じて確信の度合いが決まるのではなく、数度のアクセスにおいて判断が「ふらつく」程度（率）あるいは「迷う」程度が評定されて確信度が決定されるというものである。この仮説から予測される現象の多くは標準仮説と同様になるが、同時に Saito (1998 a) の見いだした現象を説明することができる。

妻藤 (1999 a) は、この仮説の理論的性質を検討し、判断メカニズムそのものはどのようなタイプであっても、それと、「ふらつき」仮説の組み合わせが可能であることを示し、さらに、特に多次元判断モデルの場合には、標準仮説よりも相当単純な説明原理であると結論した。

また、この仮説は、次のような含みを持っている。すなわち、もし競合する選択肢がまったくないときには、相当弱い想起であっても、相当高い確信を持つということになる。通常の再認実験のように、見かけ上選択肢がひとつの場合でも、「記憶にある」と「記憶にない」という課題の場合は、「判断」上の選択肢は2だと考えられる。それに対して、再生の場合、「思い出したものを答える」のであって、「記憶にないもの」については、何もする必要がない。すなわち、思い出せたとき、それが一つしかなく、かつそれ以外の可能性（記憶内容）に関する tip-of-the-tongue（「喉まで出かかると」現象）もないために「迷い」が発生しなければ、かなり曖昧な記憶でも、相当高い確信を持

つことになろう。知識や記憶以外の認知領域にも、この仮説が適用できるならば、「さしたる根拠もなく信じ込んでしまう」という各種の現象 (Gilovich, 1996) の基礎メカニズムである可能性も出てくるのである。この点の検討、および、上述のような Stretch and Wixted (1999) の発見したファンアウト現象が定量的に説明できるかどうかは、まだ今後の問題である。

それに加えて、標準仮定によって確信が形成される場合もあるという可能性は残されている。特に意識的に根拠を検討するような場合と、直感的な確信との相違が、そこにあるのかもしれない。これは重要な研究課題であると思われる。意識的な推論根拠の検討によって生成される「強い」情報が確信を高める（標準仮定）のか、それとも、根拠を意識的に検討することによって対立選択肢（可能性）の数が減少し、それによって「迷い」が減るために確信を強める（「ふらつき」仮説）のか、という問題である。もし後者が正しければ（つまり、すべての確信を「ふらつき」仮説で説明できるのなら）、人間の確信は、全般にわたって（研究者や実務専門家が専門的に判断するときも含めて）、かならずしも合理的なものではないことになる。要するに、根拠そのものの「強さ」ではなく対立可能性の数が問題なので、「思いつくことができる可能性の数」が少ない人は、たいした根拠もなしに、特定の選択肢に確信をもってしまう（それが正しいと信じ込む）であろう。この点は未解決であるが、いずれにせよ、長期にわたって定説に近くなっていた確信度の標準仮定に重大な問題が隠れていたことは事実であり、このメカニズムの研究は、よりアクティブになるべき分野であると結論できよう。

## References

- Barakrishnam, J.D., & Ratcliff, R. 1996 Testing models of decision making using confidence ratings in classification. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 615-633.
- Bernback, H. A. 1967 Decision processes in memory.



- Psychological Review, 74, 462–480.
- Donaldson, W., & Murdock, B. B. Jr 1968 Criterion change in continuous recognition memory. *Journal of Experimental Psychology*, 76, 325–330.
- Gilovich, T. *How we know what isn't so : The fallibility of human reason in everyday life*. 1991 The Free Press.
- Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. 1993 The regularities of recognition memory. *Psychological Review*, 100, 546–567.
- Griffin, D. & Tversky, A. 1992 The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, 24, 411–435.
- Humphreys, M. D., Bain, J. D., & Pike, R. 1989 Different ways to cue a coherent memory system : A theory for episodic, semantic, and procedural tasks. *Psychological Review*, 96, 208–233.
- Kintsch, W. 1967 Memory and decision aspects of recognition learning. *Psychological Review*, 74, 496–504.
- Maddox, W. T., & Estes, W. K. 1997 Direct and indirect stimulus–frequency effects in recognition. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 23, 539–559.
- Murdock, B. B. Jr. 1974 *Human memory : Theory and data*. Hillsdale, N. j. : Erlbaum Press.
- Murdock, B. B. Jr, & Dufty, P. O. 1972 Strength theory and recognition memory. *Journal of Experimental Psychology*, 94, 284–290.
- Norman, D. A., & Wickelgren, W. A. 1965 Short–term recognition memory for single digits and pairs of digits. *Journal of Experimental Psychology*, 70, 479–489.
- Saito, M. 1998a Fluctuations of answer and confidence rating in a general knowledge problem task : Is confidence rating a result of direct memory–relevant–output monitoring? *Japanese Psychological Research*, 40, 92–103
- 妻藤真彦 1998 b 一般知識問題における解答の記憶と確信度の変更. 日本心理学会 第62回大会発表論文集, p 853.
- 妻藤真彦 1999 a 確信度評定に関する標準仮説と対立仮説の理論的検討. 美作女子大学・同短大部紀要, 44, 8–17。
- 妻藤真彦 1999 b 一般知識問題における確信度変動分布と理論的問題. 日本心理学会 第63回大会発表論文集, p626。
- Stretch, V. & Wixted, J. T. 1999 Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 24, 1397–1410.
- 須藤昇 1993 構造的の高い刺激セットの学習における再認確信度の動的変容. 心理学研究, 63, 388–395。
- Swets, J. A., Tanner, W. P., Jr, & Birdsall, T. G. 1961 Decision processes in perception. *Psychological Review*, 68, 301–340.
- Tanner, W. P., Jr, & Swets, J. A. 1954 A decision–making theory of visual detection. *Psychological Review*, 61, 401–409.
- Wallsten, T. S. & Gonsález–Vallejo, C. 1994 Statement verification : A stochastic model of judgment and response. *Psychological Review*, 101, 490–504.