

HI評価への適性と得られる主観量に基づいた 3種の主観量測定法の比較¹

吉川 歩・西村 武 (京都工芸繊維大学 工芸学部)

Comparison among Three Psychological Scaling Methods Based on Both Theoretical and Experimental Evaluation

YOSHIKAWA, Ayumi
ayumi@dj.kit.ac.jp

NISHIMURA, Takeshi
tn@dj.kit.ac.jp

Faculty of Engineering and Design, Kyoto Institute of Technology
Matsugasaki, Sakyo-ku, Kyoto 606 JAPAN

Abstract: Both theoretical comparison among three psychological scaling methods based on theory of evaluation of Human Interface (HI) and experimental comparison among them based on subjective extent obtained from these methods are discussed. The three methods used in this study are as follows: *method of successive categories*, *fuzzy graphic rating scale* (FGRS) and *multiple-scale technique* (MUSCAT). First, the theoretical comparison makes it clear that the FGRS and the MUSCAT are suitable to evaluate HI, due to measuring subjective extent for each participant, no need of repeated measurement, and so on. Next, the experimental comparison is performed through having participants rate subjective tallness of Japanese adult men with the three methods simultaneously. The obtained results show that the three kinds of the psychological scale values for stimuli are almost equal at the level of the interval scale, and no significant difference is not found in gravity centers of psychological scale values obtained from the MUSCAT and the FGRS.

Key words : psychological scaling method, fuzzy set theory, absolute judgment, method of successive categories, fuzzy graphic rating scale, multiple-scale technique

1. まえがき

1.1 HI評価における主観量測定的重要性と特殊性

人間の主観や感性に関係した量の測定法である主観量測定法は、心理学などの人文系の分野において主として用いられてきた。しかし工学分野においても使いやすさや快適性など人間側の要素を重視するヒューマン・インタフェース(以下HI)の重要性が認識されてからは、その測定法として重要な意味を持つようになった。例えば主観量である「快適さ」と物理量である「生理指標」の相関関係を調べようとする場合には必要不可欠である[1]。本稿で扱うHIの評価とは、HIのさまざまな属性を人間の内

省報告に基づき主観的に評価することを意味する。

主観量の測定は、心理学などでもHIの評価でも、人間を測定器として測定する基本的な姿勢には変わりはないが、HIの評価には他の分野にはない特殊性もある。

反復測定が困難：心理学などの場合には実験条件の統制が容易であるが、HI評価の場合には実験条件の統制が困難な場合(つまり同じ条件で繰り返し評価したとみなし難い)が多い。これはHI評価の場合、測定対象が何らかの作業であることが多いため実験条件を揃えることが困難なことや、さらに1回の作業に長時間を要する場合には繰り返し実験することさえも困難なことなどによる。

個人毎の主観量も重要：人文系の測定では測定集団全体のコンセンサスとしての主観量の測定に重点がおかれていたが、HIの評価で扱う使いやすさや快適性などは個人差も重要な意味を持つため、各個人毎の主観量の測定も重要である。

測定が測定対象を妨害：評価対象が作業である場合、評価を測定することがその作業を中断することになることもある。

上述したような特殊性のために、現在提案されている主観量測定法の中にはHIの評価に利用し難いものも含まれている。したがって本稿の目的である主観量測定手法のHI評価への適性を明らかにすることは

原理的に適していない手法を無理に適用することを回避でき、測定対象にもっとも適した手法を示すことができる、

適性の考察を通して、新たな手法への改良点が明らかになる、

といった意義がある。これらは実験者の無駄な労力の軽減、および新しい手法の開発という意味で重要である。

1.2 代表的な主観量測定法

表1は代表的な主観量測定法[2-6]を次の三つの特徴により分類したものである[注1]。

：評価に関連するあいまいさの扱い方

(統計論 - ファジィ集合論)

：主観量の推定方法(間接推定 - 直接推定)

：判断の方法(絶対判断 - 比較判断)

表1 代表的な主観量測定法の分類
表中の分類は 統計論：あいまいさの扱い方、
ファジィ集合論：推定方法、 絶対判断：判断の方法に対応する

		絶対判断	比較判断
統計論	間接	系列範疇法	一対比較法
	直接	グラフ評定尺度法	マグニチュード推定法
ファジィ集合論	間接	多重尺度図法	ファジィ一対比較法
	直接	ファジィグラフ評定尺度法	ファジィマグニチュード推定法

注1：表1中の「ファジィマグニチュード推定法」は筆者の独断による予測・命名であり、その方法が現在存在するかは不明である。おそらく筆者の予想では基準に対して5倍くらいといった回答法を許す方法になると思われる。

まずあいまいさの扱い方について見ると、統計論の区分に分類される方法では、主観量の真値は一つの数値で表現されるという立場に立つ。これに対してファジィ集合論の区分に分類される方法では、主観量はファジィ集合として表現されるという立場に立つ。このため後者では回答の際に主観量のあいまいさ、あるいは主観量間の関係のあいまいさが考慮される。これに対し、前者ではこれらは考慮されず、それよりも反復測定の際の値の変動が重視される。

次に推定方法について見ると、直接推定に区分される手法は文字どおり評定者に主観量を直接回答させる方法である。これに対し、間接推定に区分される手法はカテゴリー尺度図を用いた評定や対比較結果をもとに心理尺度を構成し、その尺度を用いて主観量を推定する手法である。

さらに判断の方法について見ると、比較判断に区分される手法は一つの基準対象との比較、あるいは対比較のように二者の比較をもとに回答させる方法である。これに対して絶対判断に区分される手法は全対象中での個々の対象の位置を判断させる方法である。

表1に示した方法は、HIの評価に利用可能と考えられ、かつ次元属性の測定法に限定した。これらの手法は心理尺度構成法と呼ばれるものであり、間隔尺度以上の尺度の水準を満たす心理尺度値、つまり主観量を与える[注2]。

HI評価に用いる手法を選定する場合、1.1の適性と並んで重要な要因となるのが、本稿のもう一つの目的である各手法から得られる主観量の間関係である。得られた主観量を比較することの意義として次の三つを挙げることができる。

新しく開発した手法の有効性を検証できる。

異なる手法間で測定された主観量を比較検討することの妥当性が判断できる。

に関連して、とくにファジィ理論を用いた方法とファジィ理論を用いない従来法を比較することにより、主観量測定において種々のあいまいさを考慮することの有効性が検証できる。

そして上記のような視座から、SD法による結果とファジィグラフ評定尺度図法の比較[5,6]、マグニチュード推定法とメンバシップ関数例示法の比較[7]、さらにファジィグラフ評定尺度図法とファジィ範疇法の比較[8]、お

注2：対象の属性に割り当てられた数値の差まで意味を持つような物差である[2]。例えば、距離、セルシウス度などが該当する。平均値、標準偏差、相関係数などの統計量が意味を持つ。

よびファジィグラフ評定尺度図法と多重尺度図法の比較 [3]などがすでになされている。

1.3 本稿で取り上げる3種の手法

本稿で取り上げる主観量の測定法は表1中で太字で示した次の3種である。

(a)系列範疇法[2]

(b)ファジィグラフ評定尺度図法[5,6]

(c)多重尺度図法[3]

ここでこれらの手法を取り上げた理由は、(a)の系列範疇法は心理学などの人文系の分野だけでなく、HIの評価にもすでに用いられている [9-11]。したがってその適性を明らかにしておくことは重要である。(b)のファジィグラフ評定尺度図法は境界に不明瞭さを持った概念や程度などの測定法として近年よく用いられている [5,6,12,13]。また(c)の多重尺度図法は、データの収集法そのものは(a)の手法と同じであるが、尺度構成法にファジィ理論を用いたものである。現時点での応用例は少ないが、今後の応用が期待されている方法である。1.2で述べたようにこれら3種の方法のうちの一部の手法を、得られる主観量に着目して比較した研究は既に行われている。しかしながら、これらの3種の手法のHI評価への適性を明らかにし、それとともに各手法から得られる主観量、すなわち心理尺度値を比較する研究は行われていない。従ってそれらを明らかにすることは、1.1および1.2で述べたように、応用上重要である。

また表1から明らかなように取り上げた3手法は全て絶対判断型となっている。比較判断型の方法を取り上げていない理由として、

比較判断による手法の場合、対象(刺激)の増加とともに評価回数が2乗のオーダで増加し、被験者に負担となること、

系列範疇法との比較を行いたいこと、

が挙げられる。また本論文で取り上げる3種の手法は同時に測定することができるので、得られた主観量の比較を行う際に文脈効果の影響を考慮する必要はない。これらの理由から、比較判断型の手法との比較も重要ではあるが、本稿では絶対判断型に限定した。

1.4 目的および得られた結果の概要

本稿の目的は次の2点である。

3種の主観量測定法をHI評価への適性という観点から比較し、各手法のメリット、デメリットを明らかにする。

3種の手法から得られる主観量を比較し、各手法から

得られる値の間関係を明らかにする。

続く2.で3種の手法のHI評価への適性を理論的に考察し、3.では心理実験結果をもとに3種の手法から得られた心理尺度値間の比較を行っている。そして4.は全体のまとめである。

それぞれの目的に対し得られた結果の概要を示すと次のようになる。

！HI評価の特徴である、(a)反復測定条件が満たされずとは限らない、(b)各個人毎の測定結果も重要、(c)評価の回答作業ができるだけ対象に与える影響が小さいという条件に加えて、(d)日常的な判断に近いという条件を考慮すると、多重尺度図法が推奨される。一方主観量の表現力という点を考慮すると、ファジィグラフ評定尺度図法が推奨される。

3種の手法から得られる主観量は間隔尺度の水準においてほぼ同等である。また多重尺度図法とファジィグラフ評定尺度図法から得られる心理尺度値は、それらの重心位置に有意な差が認められないことが明らかになった。

2. 3種の主観量測定法のHI評価への適性の

理論的な考察

2.1 各手法の処理手順

まず3種の手法の測定原理の概略とそれらの重要な特徴を示す。また各種法を用いた尺度構成の具体例は3.2で示す。

2.1.1 系列範疇法

系列範疇法は順序尺度の性質を有する言語的なカテゴリーから構成されたカテゴリー尺度図を用いて得られた結果をもとに、心理学的連続体すなわち主観量の連続体上でのカテゴリーの心理的な距離を推定、つまり間隔尺度を構成する手法である。例えば、後出の図1(b),(c)のような評価属性と言語ヘッジから構成されたカテゴリー尺度図(図1の例では、評価属性：高い、言語ヘッジ：非常に、など)を用いて、同一刺激について反復測定を行い、その結果得られた各刺激の相対度数分布をもとに尺度構成を行う。この手法では次のような仮定を設けている [2].

カテゴリーの境界は確率的な変動を除けば明確に定まっている。

カテゴリーの選択は、カテゴリーの境界と対象の心理的な程度の比較に基づいて行われる。

同一刺激を反復測定した際に回答が複数のカテゴリーに分布する原因を確率的な変動とみなす。つまり刺激の分布は正規分布に従うとみなす。

各刺激の分布は等しい標準偏差を持つ。

これらの仮定に従えば、各刺激についてその度数分布が

正規分布となるようにカテゴリーの間隔を定め、それらを全刺激について平均することによってカテゴリーの尺度値(共通尺度値と呼ばれる)を求めることができる。つまりカテゴリーの間隔は標準正規分布の標準偏差を単位として定められる。

2.1.2 ファジィグラフ評定尺度図法

後出の図1(a)のようなグラフ尺度図、いわゆる数直線を用いて、刺激(対象)の持つ属性の主観的な程度、すなわち心理尺度値を直接回答させる方法である。この方法は心理尺度値の境界の不明瞭さ(以後、ベグネス)をファジィ集合を通して扱うことができる^[注3]。この手法はメンバシップ関数の同定法という観点からは、該当要素応答法に分類される^[14]。

Heskethらの提案した方法では「最もよく表している点」と「その点の変動する可能性のある範囲」を回答させ、三角型メンバシップ関数を構成し、その心理尺度値を表現している^[5,6]。また吉川・西村^[15]はこの方法を改良し、「最もよく表していると思う範囲」と「表しているとみなせる範囲」の二つの区間を回答させ、台形型メンバシップ関数を構成し、その心理尺度値を表現している。本稿では、後者の方法を採用した。

2.1.3 多重尺度図法

多重尺度図法のもととなるファジィ範疇法は、カテゴリーおよび対象の主観量の両者のベグネスを考慮して、カテゴリー尺度図から得られた評定結果に対する心理尺度値をファジィ集合として推定する方法である。しかし系列範疇法との類似点は、カテゴリー尺度図を用いる点のみであり、単に系列範疇法をファジィ化した方法ではない。

ファジィ範疇法のもととなる評定判断過程のモデルでは、以下のような仮定が設けられている^[15]。

刺激の持つ属性の主観的な程度は「刺激が属性を持つ」という命題の真理値と等価である。

この真理値とカテゴリーと同じ言語ヘッジを持つ言語真理値の一致の程度をもとにカテゴリーが選択される。

属性の主観的な程度、それと等価な真理値、言語真理値および真理値と言語真理値の一致の程度はファジィ集合として表現される。

一致の程度の算出は命題の真理値を言語真理値で真理値限定することにより行われる。

注3：広義にはメンバシップ関数例示法のような尺度図上の1点だけを回答させる方法や区間推定法のような所属度1の範囲だけを回答させる方法も、この手法の一種とみなすことができる。これらの手法の解説は文献^[21]に詳しい。

さらにここで

真理値限定から得られた一致の程度をunitary true^[注4]と仮定する

ことにより、カテゴリーと同じ言語ヘッジを持つ言語真理値の形状がカテゴリーの心理尺度値の推定値として与えられる(例えば、「非常に高い」の心理尺度値は「非常に真」の形状として与えられる)^[8]。これより、この言語真理値の形状をファジィグラフ評定尺度図法を用いて同定することがカテゴリーの尺度構成に相当し、そしてこの形状を用いることで刺激の心理尺度値が与えられる。主観的な背の高さの場合を例にとると、ある人が「非常に背が高い」と評定されたら、その人の背の高さの心理尺度値は「非常に真」のメンバシップ関数により表されることになる。

ただし、カテゴリー尺度図が離散的であるために、得られる刺激の心理尺度値は言語カテゴリーの形状によって規定される離散的な値しか取れないことになる。この欠点を緩和するために提案されたのが、多重尺度図法である^[3]。この方法では、同一刺激を異なるカテゴリーから構成された同一属性に関する複数の尺度図を用いて測定し、各尺度図から得られた結果を合成することにより上記の欠点を緩和するものである。吉川らは心理実験を行い、合成に用いる尺度図は2本、合成法はmax合成後に凸化演算を行う方法が推奨されることを明らかにしている^[3,16]^[注5]。本稿でも2本の尺度図から得られた結果をmax合成する方法を用いている。この手法はファジィグラフ評定尺度図法と同様、1回の評定からベグネスを考慮して刺激の心理尺度値を推定することが可能である。

2.2 HI評価に対するメリット、デメリット

表2は3種の主観量測定法の特徴を比較したものである^[17]。6種の比較項目のうち上の3項目が1.1で触れたHI評価の特殊性を強く反映している項目であり、残りの3項目は主観量測定一般に共通するとみなせる項目である。

まず上の3項目について見てみると、多重尺度図法およびファジィグラフ評定尺度図法とも評価は高く、これらの方法がHI評価に適していることが見て取れる。両者の差異は、測定対象への妨害性に現れている。これは多重尺度図法の場合カテゴリー尺度図を使わず口頭で評価を回答させるといった工夫を行うことにより作業への影響

注4：数値真理値 $[0,1]$ を定義域、所属度 $[0,1]$ を値域とするとき、傾き1、切片0(つまり、数値真理値=所属度)となるtype-IIファジィ集合。

注5：ファジィ集合A,Bのmax合成+凸化は「AからBまで」を表す集合を求めていることに相当する^[22]。

表2 3種の主観量測定法の特徴の比較

項目	方法	
	多重尺度図法+ファジィ範疇法	
	ファジィグラフ評定尺度図法	カテゴリー尺度図+系列範疇法
一回の評定から心理尺度値が得られる	×	
評定者ごとに心理尺度値が求められる		
評定が他の作業の妨げとならない		
得られた心理尺度値の表現力が大きい		
日常的な判断に近い形で回答できる		
実時間で心理尺度値が求められる	×	

○ : 全く問題なし, ◐ : ほぼ問題なし,
 ◑ : やや問題あり, × : 不可能

を軽減できるのに対し、ファジィグラフ評定尺度図法の場合は必ずグラフ尺度図への評価の記入が必要なことによる。また表より系列範疇法はHI評価に適しているとは言い難いことがわかる。

次に下の3項目について見てみると、直接推定法であるファジィグラフ評定尺度図法および刺激の評定に先立ちカテゴリーの心理尺度構成が行えるファジィ範疇法は実時間で刺激の心理尺度値が求められるが、系列範疇法はデータの集積が必要なため実時間の尺度構成は不可能である。また多重尺度図法は「非常に」などの言語ヘッジを通して属性の程度を回答できるので、グラフ尺度図上に程度を直接回答するファジィグラフ評定尺度図法に比べて、より日常的な程度の判断に近い形で回答できる。このため初心者にも回答しやすいと考えられる。しかしその反面、カテゴリー尺度図が離散的であるために、ファジィグラフ評定尺度図法に比べて刺激の心理尺度値を表現する際にその自由度が低くなるという欠点もある。このため尺度図を構成するカテゴリーの選択の仕方が、心理尺度値の表現力を左右することは言うまでもない。

これらの考察をまとめると次のようになる。

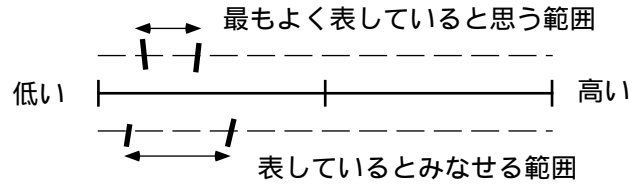
ファジィグラフ評定尺度図法、多重尺度図法ともにHIの評価に適している。そして

- ・ 作業への低妨害性、日常的な判断に近い点を重視すれば多重尺度図法
- ・ 心理尺度値の表現力を重視すればファジィグラフ評定尺度図法

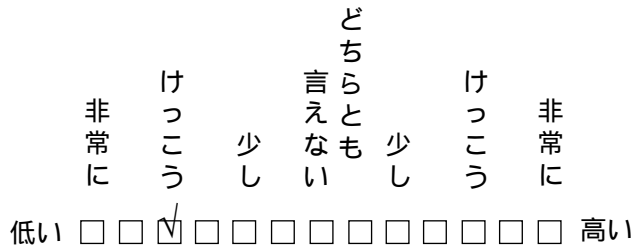
が推奨される。これらの方法は測定対象、回答者(被験者)に応じて適宜使い分ければよい。

系列範疇法はHIの評価に適しているとは言い難い。しかし集団全体の共通評価だけが必要で、被験者数が十

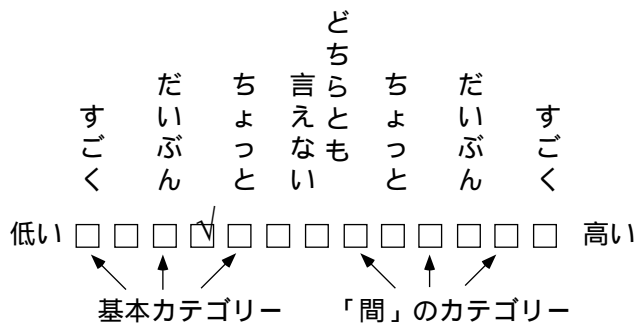
140 cm



(a) グラフ評定尺度図



(b) カテゴリー尺度図1



(c) カテゴリー尺度図2

図1 実験に用いた回答用紙の例とその回答法

分多い場合にはその有効性は失われない。

3. 3種の主観量測定法から得られる主観量の 実験による比較

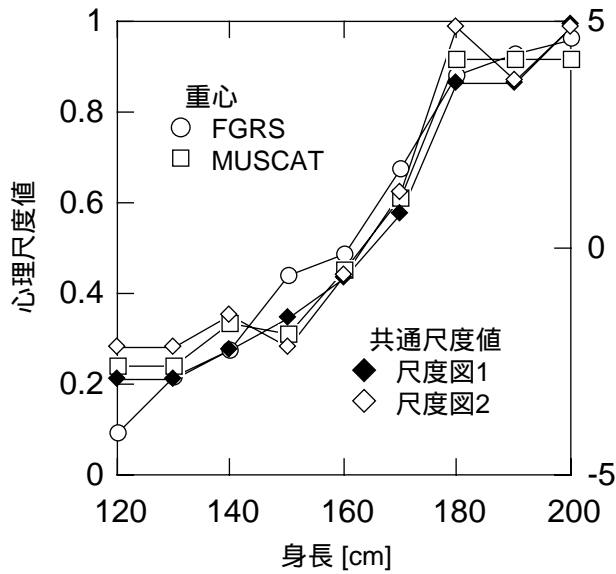
3.1 実験条件および手順

被験者：京都工芸繊維大学の学生および教職員であった。彼らは日本語を母国語としており実験内容の理解については問題がない。またすべての被験者はボランティアとして参加した。記入洩れや回答方法に誤りのある被験者を除いたデータ解析可能な被験者は54名であった。

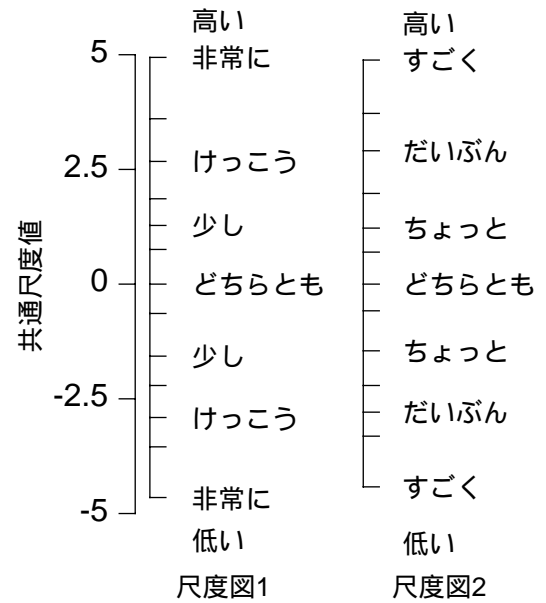
刺激：日本人男性の数値で表された身長とした。120cmから200cmの間を10cm刻みに分割した9種を用いた。

回答用紙：主観的な高さの評定用と言語真理値の同定用の2種を用いた。

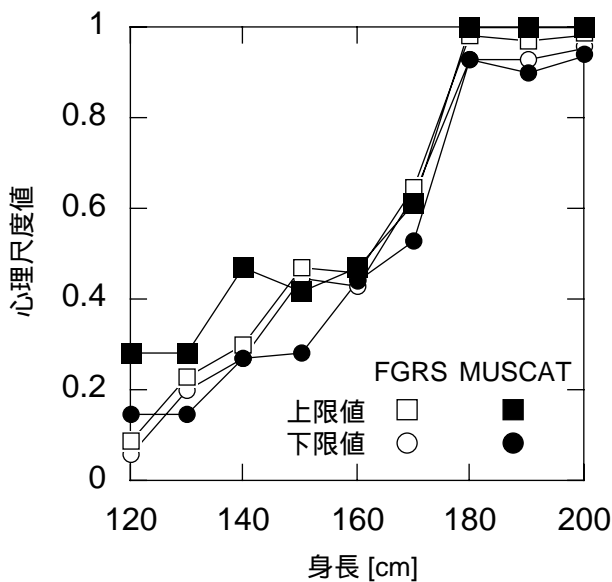
主観的な高さの評定用の回答用紙には、1ページに図1に示すような、2つのカテゴリー評定尺度図と1つのグラフ評定尺度図が印刷されていた。カテゴリー尺度図は系列範疇法と多重尺度図法で用いるデータの収集に利用され、一方グラフ尺度図は



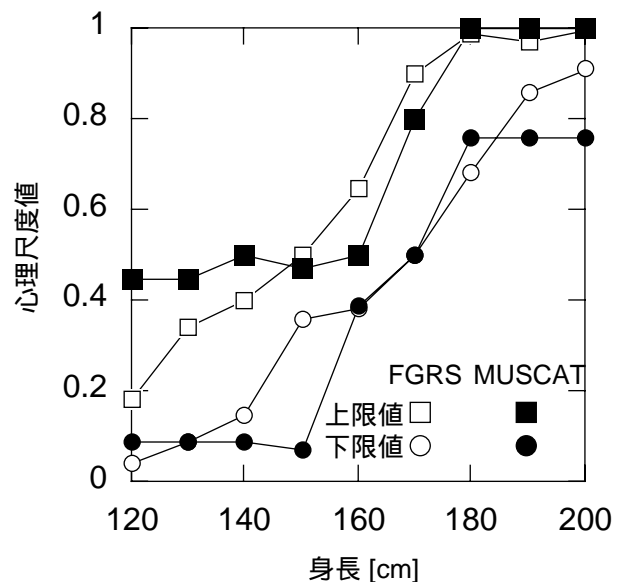
(a)心理尺度値の重心と共通尺度値



(b)尺度図1および2の各カテゴリーに関する共通尺度値



(c)心理尺度値の1レベル集合の上限と下限



(d)心理尺度値の台集合の上限と下限

図2 1名の被験者の実験結果：図中FGRSはファジィグラフ評定尺度図法による心理尺度値，MUSCATは多重尺度図法による心理尺度値を意味する。

ファジィグラフ評定尺度図法による尺度構成に利用された。

グラフ尺度図は長さ10cmで、その両端と中央にアンカー(縦棒)が設けられていた。被験者は左上の数値的な身長に該当する自分の身の回りに居る日本人成人男性を想像するように求められ、その人の背の高さ-低さの程度を「最もよく表していると思う範囲」と「表しているとみなせる範囲」をそれぞれの区間の両端の2点を記入することにより回答した(図1(a)参照)。

また各カテゴリー評定尺度図は、例えば「非常に高い」のような言語ヘッジにより表された7つの基本カテゴリーとそれらの基本カテゴリーの“間”を表すカテゴリー6つの計13カテゴリーから構成されていた。被験者は、与えられた数値的身長を最も

よく表しているカテゴリーを各尺度図それぞれ一つずつ選択することによって回答した(図1(b),(c)参照)。

言語真理値の同定は、先の主観的な高さの評定と同様、グラフ尺度図を用いたファジィグラフ評定尺度図法により行われた。同定する言語真理値は、先のカテゴリー尺度図の基本カテゴリーに用いられていた言語ヘッジと真あるいは偽を組合わせたもの12種類(例えば「非常に真」と「真でも偽でもない」の計13種であった。これらを1ページ当り5種、3ページにわたって印刷した。

表紙に引き続き、先の2種類の回答用紙を主観的な背の高さの評定10ページ、言語真理値の同定用の4ページの順で綴じ1冊の

回答用紙とした．なお上記のページ数には，2種類の回答用紙の先頭にある各課題の回答法の教示のページを含んでいる．全被験者とも同じ冊子を用いて実験を行った．

実験手順：各被験者に回答用紙を配布して回答してもらい，後日回収した．人と相談せず必ず一人で回答する旨の教示を回答用紙の表紙に印刷するとともに配布時に口頭で伝えた．

3.2 実験結果

3.2.1 各手法から得られた心理尺度値の一例

系列範疇法：図2(b)は全被験者の評定結果を用いて算出したカテゴリーの共通尺度値である．意味上の原点である「どちらとも言えない」を0とおいた．図から「どちらとも言えない」付近の間隔が狭く，末端カテゴリーである「非常に」あるいは「すごく」とその一つ内側のカテゴリーとの間隔が広くなるという系列範疇法から得られる結果によく見られる傾向が現れている．また図2(a)の菱形で示したデータは，このカテゴリーの共通尺度値と評定結果をもとに1名の被験者の刺激に関する心理尺度値を表したものである．

次にこのカテゴリーの共通尺度値と各刺激に対する評定結果の分布を用いて，その分布の正規性および等分散性の検定を行った．その結果をまとめたものが表3である．ここではギルホードに倣い，共通尺度値をもとに計算した期待累積比率と実測累積比率の食い違い量を χ^2 検定するMostellerの方法を用いた[2]^[注6]．表から実測比率と期待比率から算出した χ^2 値は，尺度図1，2それぞれ46.4，43.9(自由度96)となり，有意水準を0.995($\chi^2=64.06$)としても「両者に差はない」という仮説は棄却されなかった．一方，等分散に関する χ^2 値は，尺度図1，2それぞれ66.2，49.2(自由度8)となり，0.01の点($\chi^2=20.09$)を超えて有意である．これらより各刺激に対する分布の形状は正規型で

表3 系列範疇法の仮定に関する検定結果

検定項目	各尺度図の χ^2		df	100 α %点
	1	2		
正規性	46.4	43.9	96	$\chi^2_{0.995}(96) = 64.6$
等分散性	66.2	49.2	8	$\chi^2_{0.01}(8) = 20.1$

注6：ギルホード[2]にはカテゴリー境界の尺度構成をした場合の計算例だけが示されている．本稿ではカテゴリー重心の尺度構成を行っているため，実測比率を重心に対応させるために，あるカテゴリー以下と判断された比率を，そのカテゴリーより下と判断された比率+そのカテゴリーと判断された比率 $\times 0.5$ として補正した．

あるが，それらの間の分散は等しくないことがわかる．つまり厳密には系列範疇法により尺度構成するための前提となる仮定が満たされていないことを意味する．このような身長という比較的簡単な刺激を用いた場合でも厳密には仮定が満たされていないことから，HIの評価で系列範疇法の仮定を満たすように実験を行うことが容易でないことがわかる．

ファジィグラフ評定尺度図法：グラフ尺度図上に回答された二つの区間からファジィ集合を構成する手順は次のようである．グラフ尺度図上に被験者により回答された2つの範囲を規定する4つの境界を物差を用いて1[mm]単位で読み取った．それらの値をグラフ尺度図の左端を0，右端を1として数値化した．二つの範囲のうち「もっともよく表していると思う範囲」を1レベル集合，「表しているとみなせる範囲」を台集合とし，そしてその間の尺度値は線形補間を行い台形型のメンバシップ関数を構成した^[注7]．このようにして得られた結果の一例を図2(c),(d)に示した．

図2(c)から，ここに示した被験者は「もっともよく表していると思う範囲」を幅を持った区間として回答していることがわかる．また全被験者の1レベル集合幅の平均値は4.6[mm]，その標準偏差は2.8[mm]であった．この値を用いて，この標本平均が母平均0に一致するという仮説をt-検定により検定した結果，t値は36.24となり自由度485の片側0.05%点3.29を超え仮説は棄却された．このことはHeskethらが用いているような「最もよく表している範囲」を最初から一点として回答させる方法[5,6]が必ずしも適切でないことを示している．

多重尺度図法：上述のファジィグラフ評定尺度図法における刺激のメンバシップ関数の場合と同様の手順により言語真理値のメンバシップ関数を構成した．2.1.3で述べたように，これらの言語真理値がそれぞれ対応するカテゴリーの心理尺度値の推定値となる．なお文献3,8,17では例えば，「非常に高い」と「非常に低い」が0.5を対称軸として線対称になるように，「非常に低い」を「非常に真」のtype-IIの否定として求めていた^[注8]．しかしこれまでの研究から例えば「非常に真」と「非常に偽」が0.5を対称軸としてほぼ線対称となることが得られているので，本稿でも文献16と同様「非常に偽」を「非常に低い」の推定値として用いた．また基本カテゴリーの間を表す

注7：台集合という用語は，所属度が0でない要素の区間を指す場合と考慮している全体集合を指す場合があるが，本稿では前者の意味で用いている．後者に対しては全体集合という語を用いる．

注8： $f'(x) = f(1-x)$ として定義される． $f'(x) = 1 - f(x)$ が通常type - Iの否定．ただし $x \in [0,1]$ ．

カテゴリーは両側の基本カテゴリーの*Between*集合[18]として算出した．次に被験者の評定結果に対応する各尺度図のカテゴリーの心理尺度値をmax合成することにより，刺激の心理尺度値を求めた．このようにして得られた結果の一例を図2(c),(d)に示した．

3.2.2 手法間における心理尺度値の比較

(a) 相関係数

1.2で触れたように，3種の手法から得られる心理尺度値は少なくとも間隔尺度の水準を満足する．そこで各手法から得られた心理尺度値間の関係を明らかにするために相関係数を算出した．ところで，ファジィグラフ評定尺度図法と多重尺度図法から得られる心理尺度値はファジィ集合であるため，通常の数値データに適用される方法で相関係数を求めることができない．そこで文献8と同様，ファジィ集合の重心を求めることにより非ファジィ化した．図2(a)は1名の被験者の二つの重心と系列範疇法の心理尺度値である．これらを用いて推定法間の相関係数を算出し，3種の手法間の比較を行った．

表4は3種の測定法から得られた心理尺度値をもとに算出した相関係数である．表より被験者毎に求めた相関係数の全被験者についての平均値は，いずれの手法の組合せに対しても0.98以上となった．また各被験者ごとの相関係数の中で最小となった値でも0.946であった．この値は自由度7(=刺激数9-2)の相関係数の両側0.1%点0.898を超えて有意である．図2に示した結果はこの被験者に対するものである．さらに全被験者の値を用いて相関係数を求めた結果，もっとも小さな値になったファジィグラフ評定尺度図法と系列範疇法の尺度図2の組合せでも0.98以上となった．これらの結果から考えると，3つの手法から得られた心理尺度値は，間隔尺度の水準において同等であると結論付けることができる．

(b) 一致度，類似度および形状，位置の指標

文献3,16と同様，ファジィグラフ評定尺度図法と多重尺

表4 3種の測定法から得られた心理尺度値の間の相関係数

手法の組合せ	ファジィグラフ評定尺度図法		多重尺度図法		
	多重尺度図法	系列範疇法		系列範疇法	
		尺度図1	尺度図2	尺度図1	尺度図2
被験者間平均値	0.988	0.988	0.985	0.993	0.991
標準偏差	0.009	0.009	0.011	0.007	0.006
最小値	0.963	0.958	0.946	0.957	0.971
最大値	0.998	0.998	0.997	0.999	1.000
全被験者	0.980	0.983	0.980	0.989	0.987

度図法から得られたファジィ心理尺度値の関係を一致度と類似度の二つの指標を用いて比較した．相関係数が間隔尺度の水準での比較を行っていたのに対し，これらの二つの指標は比率尺度の水準での比較を行っていることに相当する．一致度はその二つのファジィ集合から求められた論理積集合の所属度の最大値として与えられる．また類似度はその二つのファジィ集合から論理積集合と論理和集合の濃度を求めて，論理和集合の濃度に対する論理積集合の濃度の比として与えられる^[注9]．前者は二つのファジィ集合が最もよく一致している要素での所属度を与えるのに対して，後者は二つのファジィ集合の位置および形状の厳密な一致の程度を与える．両指標とも0から1の間の値を取り，1に近いほど一致の程度がよい．

表5に示したように，上記の二つの指標の全被験者，全刺激に関する平均値(標準偏差)は，一致度が0.871(SD:0.274)，類似度が0.363(SD:0.240)であった．母平均と標本平均の差に関するt検定を用いて，一致度の標本平均が1に等しいという仮説を検定したところ，t値は0.47(自由度485)となり，有意水準を片側0.1(t=1.282)にとっても仮説は棄却されなかった．これに対し，同様の検定を類似度について行ったところ，t値は2.65となり有意水準0.01(片側，t=2.326)で仮説は棄却された．これらの結果からわかることは，二つの手法から得られた心理尺度値は位置と形状を考慮した厳密な意味での一致は得られていないものの，二つの心理尺度値の間に所属度が1となる要素は存在する．しかしながら類似度の低下は位置のずれと形状の相違の二つにより引き起こされるため，この結果だけではそれぞれの要因の影響の度合を推定することはできない．

そこで形状の差異と位置のずれの影響を推定するために次のような指標を算出した．位置のずれを評価する指標として，二つのファジィ集合の重心間の距離を1から引

表5 2つの手法の心理尺度値の関係を表す4種の指標の検定結果

指標	平均値(SD)	t値	df	100α%点(片側)
一致度	0.871(0.274)	0.47	485	$t_{0.10}(485) = 1.282$
類似度	0.363(0.240)	2.65	485	$t_{0.05}(485) = 1.645$
形状	0.532(0.227)	2.06	485	
位置	0.944(0.049)	1.14	485	$t_{0.01}(485) = 2.326$

いたものを用いた．重心間距離の絶対値を1から引いた理

注9：濃度はファジィ集合が連続の場合には，そのメンバシップ関数の面積として与えられる．

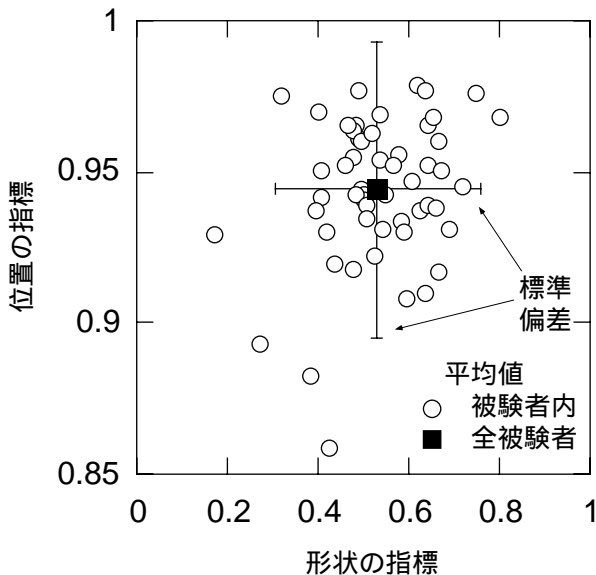


図3 平均値による形状の指標と位置の指標の関係

由は他の指標と同様に1に近づくほど位置に関してよい一致を示すようにするためである[注10]。一方形状を評価する指標として、二つのファジィ集合の重心を一致させたときの類似度を用いた。ファジィ集合の重心を一致させることにより、位置のずれの影響を完全ではないものの、ある程度除外することができる。これらの指標をそれぞれ形状と位置の指標と呼ぶ。

図3は形状の指標と位置の指標の関係を表したものである。被験者内平均値で表示した。全被験者に関する形状および位置の指標の平均値はそれぞれ0.532(SD:0.227)、0.944(SD:0.049)であった(図中の、表5参照)。一致度、類似度の場合と同様に、母平均と標本平均の差に関するt-検定により標本平均が1と等しいという仮説の検定を行ったところ、形状、位置に対してそれぞれ、t値2.062、1.143となった。前者は片側0.05の点(1.645)を超えて有意となり、後者は片側0.10の点(1.282)で有意と判断されなかった。ファジィグラフ評定尺度図法から得られる心理尺度値は位置、形状とも完全に制約がないのに対し、多重尺度図法から得られるそれは基本的にカテゴリーの心理尺度値の形状に依存する離散量である。このことより両者の形状の一致は望めない。にもかかわらず位置の指標として表されている両者の重心間の位置に有意な差が認められないことは特筆に値する。また評価の際の回答の自信の程度を考慮すること[19]やファジィグラフ評定尺度図法の改良を行うこと[20]によって、さらに改善される

注10：ここでは全体集合が[0,1]であるため単に1から引くだけでよいが、任意の大きさの全体集合の場合には1から引く前に全体集合上での最大距離で重心間距離を正規化する必要がある。

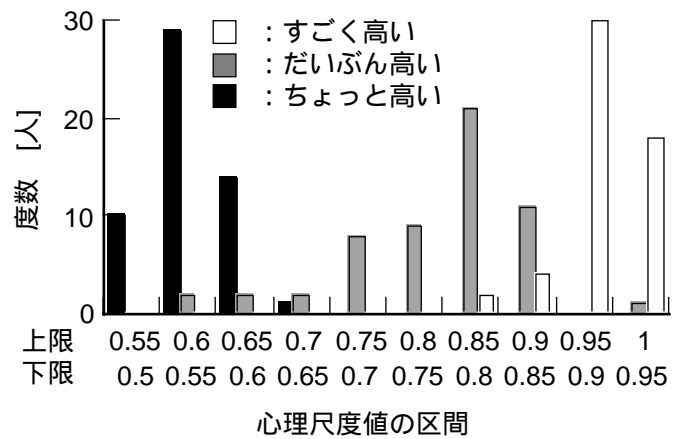


図4 3種のカテゴリーの心理尺度値の重心の分布

可能性がある。

3.3 考察

3種の測定法から得られた心理尺度値の間で有意な正の相関が認められたことは、従来系列範疇法を用いて測定されていた対象を他の二つの手法により測定することが可能であること示している。さらに2.2のHI評価への適性の比較結果を併せて考えると、評価者全体の共通評価のみが重要で、被験者数が多く、しかも反復測定とみなすことができるといった条件が整っているという特別な場合を除けば、ファジィグラフ評定尺度図法あるいは多重尺度図法を用いることが推奨される。ただし3.2.1で触れたように、系列範疇法の要求する仮定を厳密に満たすことは困難が伴うと予想される。

最後に評価カテゴリーの個人差について述べる。1.1で触れたようにHIの評価では個人差を測定することも重要である。ところが、系列範疇法の場合には、たとえ被験者間でカテゴリーの表す程度に差があっても、各カテゴリーは単一の値により表現されてしまう。しかし今回の実験でファジィ範疇法から得られた心理尺度値から被験者ごとに各基本カテゴリーの重心を計算し、その分布を求めたところ図4に示すような結果が得られた。これより無視することのできない個人差が存在することが明らかである。また「ちょっと-だいぶ」では0.55~0.7、「だいぶ-すごく」では0.8~0.9のそれぞれの区間で、被験者の間でカテゴリーの表す大小関係が逆転している例も見られる。これらの結果は系列範疇法を用いる際には、慎重な取り扱いが必要なことを示唆している。

4. まとめ

本稿では絶対判断型の主観量測定法3種について、HI評価に対する適性の理論的な考察、および各手法から得られる主観量の比較を行った。以下のような結果が得られた。

多重尺度図法およびファジィグラフ評定尺度図法がHI評価に適している。作業への低妨害性、日常的な判断に近い点を重視すれば前者、得られる主観量の表現力を重視すれば後者となる。

3種の手法から得られる主観量は間隔尺度の水準で同等であることが明らかになった。また多重尺度図法とファジィグラフ評定尺度図法から得られる心理尺度値は、それらの重心位置に有意な差が認められないことが明らかになった。

謝辞：心理実験のデータ収集の一部を担当していただいたMs. Moriani Mohamed に深謝します。

参考文献

- [1] 寺下ほか：主観的感情状態の因子分析による多面的評価，第8回ヒューマン・インタフェース・シンポジウム講演論文集，217-222 (1992)
- [2] ギルホード(秋重監訳)：精神測定法，培風館 (1959)
- [3] 吉川・西村：多重尺度図法とBetween集合によるファジィ範ちゅう法の改良，*信学論D-II*, J77-D-II, 1, 154-161 (1994)
- [4] 長沢：ファジィ構造モデルの官能検査への適用，第6回ファジィシステムシンポジウム講演論文集，349-352 (1990)
- [5] Hesketh et al. : Practical applications and psychometric evaluation of a computerized fuzzy graphic rating scale, In Zetenyi (Ed.): *Fuzzy set in psychology*, 425-454, North-Holland (1988)
- [6] Hesketh et al. : An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences, *Int. J. of Man-Machine Studies*, 29, 21-35 (1988)
- [7] Yamashita & Yamashita : Exemplifying and comparing fuzzy ratings and magnitude estimations, *Psychologica*, 35, 4, 240-248 (1992)
- [8] 吉川・西村：ファジィ範疇法による心理尺度構成とその実験的検証，*ファジィ誌*, 5, 4, 719-731 (1993)
- [9] 中山：心理尺度構成法とその応用，1987ヒューマン・インタフェース講習会資料，77-130 (1987)
- [10] 成田：主観評価で用いる日本語評価用語の知覚間隔の検証，*信学論D-II*, J75-D-II, 11, 1968-1974 (1992)
- [11] 西村ほか：VDT陽画表示画面の好ましい白背景色，*テレビ誌*, 42, 12, 1351-1357 (1988)
- [12] 竹内：ファジィ評定法による程度表現用語の意味計測，*計量国語学*, 17, 8, 365-376 (1991)
- [13] 竹村：言語的確率表現語の心理学的研究，第6回ファジィシステムシンポジウム講演論文集，335-338 (1990)
- [14] 中村：ファジィ理論と人文・社会科学(第5章ファジィ理論と行動科学)，189-240，日刊工業新聞社 (1994)
- [15] 吉川・西村：評定判断過程の新モデルとその実験的検証，*ファジィ誌*, 3, 2, 366-371 (1991)
- [16] Yoshikawa & Nishimura : Comparison among methods for compounding psychological scale values in the Multiple-scale Technique, *IEICE Trans. on Funda.*, E77-A, 7, 1202-1205 (1994)
- [17] 吉川：言語表現された主観的程度の定量化に関する研究，京都工芸繊維大学博士論文 (1992)

- [18] 吉川：Between集合の数学的性質，*ファジィ誌*, 4, 1, 150-159 (1992)
- [19] 吉川・西村：評価の自信の程度を利用したファジィ範疇法と多重尺度図法の改良，第4回ノンエンジニアリングファジィ・ワークショップ講演論文集，60-63 (1994)
- [20] 吉川ほか：計算機の援用によるメンバシップ関数同定法の改良，第4回インテリジェント・システム・シンポジウム講演論文集，7-12 (1994)
- [21] 山下：ファジィ - 心理学への展開 - ，垣内出版 (1992)
- [22] シュマッカー(鬼沢訳)：ファジィ集合，啓学出版 (1990)

[著者紹介]

吉川 歩(よしかわ あゆみ)

平成4年京都工芸繊維大学大学院工芸科学研究科博士後期課程単位修得退学。同年4月京都工芸繊維大学工芸学部電子情報工学科助手。人間の行う主観的な情報処理過程や人間によって処理された主観情報を解析・応用する「主観情報処理」に関する研究に従事。その一環としてファジィ理論を用いた心理尺度構成法の開発，ファジィ集合の同定法の改良およびファジィ集合を用いた「間(あいだ)」の表現方法などに関する研究を行っている。京都工芸繊維大学博士(学術)。平成6年電気関係学会関西支部連合大会奨励賞受賞。日本ファジィ学会，計測自動制御学会，電子情報通信学会，日本行動計量学会，システム制御情報学会，IEEEなどの会員。

西村 武(にしむら たけし)

昭和33年京都大学工学部電気工学科卒業。同年(株)日立製作所入社，中央研究所勤務。昭和43年京都工芸繊維大学に転じ，工芸学部助教授，工業短期大学部教授を経て，昭和63年工芸学部電子情報工学科教授，平成7年度より工芸学部長。この間，画質評価の研究，ヒューマン・インタフェースの評価の研究などに従事。昭和50年テレビジョン学会丹羽高柳賞論文部門受賞。平成6年度照明学会関西支部支部長。工学博士。計測自動制御学会，テレビジョン学会，照明学会，日本照明委員会，電子情報通信学会，日本色彩学会，日本人間工学会，日本社会心理学会各会員。

(1994年11月27日受付 1995年1月23日改訂)