

交叉文化的消費者研究における方法論的問題

阿 部 周 造

はじめに

本稿は消費者行動の交叉文化的研究のためのデータ収集と分析方法を考察するものである。マーケティング論の分野において、消費者行動の国際比較や交叉文化的研究の必要性が認識されたのは最近のことではない。そしてそうした研究の必要性と共に認識されてきたのが研究の困難性である (Berent 1975, Mayor 1980)。しかし、その研究の困難性についての認識の割には国際間や異文化間の比較可能性についての十分な検討がなされてきているとはいえないのが実情である。この分野における知識の積み重ねが、単なる研究数の増加だけではなく、方法論的に比較可能性の検討をふまえた研究の積み重ねによってはじめて本格的なものとなるものであることを鑑みると、比較可能性について方法論的基盤を整理確立しておくことは焦眉の課題であると考えられる。

かかる問題意識に立って、筆者は本誌に発表された「消費者行動の国際比較：その予備的考察」(1984)において国際比較の方法を論じている。そこでは用いられる統計的分析方法に応じてどこまで意味のある比較考察が可能かという視点から考察を行なった。そして複数の研究者によって、それぞれ独自の研究が進められるという状況の中での比較可能性を考えるかぎり、明示的な理論モデルを有し、かつ構成概念妥当性の検討のなされるアプローチが最も比較可能性が高いことが明らかとなった。本稿において

も、構成概念妥当性の検討の重要性は強調されるポイントとなる。研究において測定せんとした構成概念が測定されているか否かの検討は比較可能性を問う場合にも欠くことのできぬことだからである。しかし本稿は前稿と違って複数の独立した研究間の比較可能性ではなく、単一の研究における複数国間あるいは文化横断的な比較可能性を論じようとするものである。研究者が複数の文化あるいは国にまたがった消費者行動の比較研究を行なおうとする場合、そこにどのような問題があり、どのような点に注意しなければならないのかということである。

本稿はわれわれの関心事とする消費者行動の研究が心理的側面の大きい現象分野である事に鑑み交叉文化心理学の分野における研究に依拠しつつ、消費者行動の比較可能性について考察を行なうことにする。そしてファースト・フード・レストランについての日本と米国における学生のサンプルを例として、現在最もフォーマルな形での比較の方法であると考えられる LISREL のマルチ・サンプル分析を用いて比較可能性が具体的にどのように確認されるのかを例示することにしよう。尚、本来は交叉文化的 (cross-cultural) 研究の方がより一般的な用語なのであろう (Dant & Barnes Jr. 1988) が、ここではわが国が比較的単一の文化圏よりなることから、よりなじみやすい複数の国にまたがった (cross-national) あるいは国際 (international) 比較研究という用語を用いることにする。

消費者行動の比較可能性

複数の国にまたがった消費者行動の調査結果が比較される場合、どこまで比較可能なものとして考えられるのだろうか。言い換えれば消費者行動の国際比較は、国内における比較研究とくらべてどこに特別の注意を払わなければならないのであろうか。国内あるいは同一文化内において消費者行動の比較（例えば男女間の比較）がなされる場合には、得られたデータの比較可能性にとって深刻な問題が生ずる可能性は相対的に少ない。標本誤差や測定誤差について困難な問題はあっても研究者にとって思いがけない要因によって比較不能な状況が生まれてくることは稀である。それに対して、国際比較の場合には国内比較研究における問題に追加される形で文化が異なるために生じてくる比較上の問題を考えねばならなくなる。それは比較しようとしている事がどこまで比較可能かを検討しなければならないということであって、たとえば文化差を調べようという研究の前にある程度文化差を知っておかないとそれが旨くいかないといった複雑性を含んでいる。

まず異文化間の比較可能性について、交叉文化心理学で指摘されている検討点に添って考察を進めることにしよう。それらは比較されるべき概念の意味内容、その操作化、及び測定における相等性であり、そしてデータ収集方法の相等性である。以下、より概念的レベルのものから具体的レベルのものへと順を追って相等性の内容を検討してゆこう。

概念的相等性 (conceptual equivalence)

比較可能性においてまず検討されなければならないことは複数の国で用いられる概念が等しいものか否かということである。表面的には同一のものと見えるものであっても、国によってその有する意味がかなり違っているということがある。たとえば一つの料理であっても一方の国にとっては郷土の料理であるのに対して他方

の国ではエキゾチックな料理とみなされるならばその料理は違った意味を有していることになる。概念的相等性は機能的相等性 (functional equivalence) とも密接な関係を持っていると言われる (Hui & Triandis 1985)。機能的相等性は比較される概念が複数の国において他の概念や変数と同じ関係を有しているかということである。ある国における自転車の購買がスポーツ目的でなされるのに対し、別の国では輸送手段として購買される場合には同じく自転車の購買という行動ではあっても機能的相等性が乏しいということになる。こうした概念的差異および機能的差異は関心とする概念の意味範囲やかかわりを持つ範囲の差異とも考えられるから、カテゴリー相等性 (category equivalence) として問題にされることもある (Douglas & Craig 1983)。国際比較研究において概念的相等性と機能的相等性の検討は欠かすことのできないものである。場合によっては一つの概念のこうした相等性を明らかにすることが比較研究の中心的目的となることも考えられる。収集されたデータを単純に数値として比較しているだけでは見過ごされてしまう質的な差異の問題こそまさにこの概念的相等性と機能的相等性にかかわる問題であるといえよう。

しかしながら、この二つの相等性は比較可能性の検討において克服し難い事柄であるのかというところではない。研究者が複数の国における消費者行動の実情に通じていることによって、あるいは、それぞれの国の研究者による共同研究のかたちがとられることによってこの相等性の問題を押えることは可能である。概念的相等性は高い事が望ましいことは言うまでもないが、複数の国で用いられる概念に違いがあったとしても、どのような違いがあるかについてはっきりと押えられているときは、それだけ比較可能性を回復できる点に注意せねばならない。概念的及び機能的にどこまで相等性があるのかを掴まれない比較研究が比較可能性についての問題を最も多く孕んでいるのである。

概念の操作化における相等性 (equivalence in concept operationalization)

関心事とする概念の相等性が確かめられると次のステップはその概念の測定のための操作化にあたってどこまで比較対象国間で相等性が認められるかということである。概念の意味や機能上の差異がある場合にはもちろんのこと概念的相等性が高い場合でも測定にあたって異なった操作化を行わなければならないことはあり得る。たとえば、職場において贈り物をするということが、人間関係の円滑化や、常日頃お世話になっていることへの感謝の表明という目的でなされることは共通であっても、贈り物が相手の家を訪問してなされるのか職場でなされるのか、あるいはわが国のように店頭からの配送というかたちでなされるのかによって行動の測定の仕方は異なってくる。又、贈答行動のなされる時も、お歳暮とクリスマスのようにある時期に集中してなされるのか誕生日のような個人的な時に合わせて行なわれるのかと異なっている。そうした場合どちらか一つの仕方によってのみ贈答がなされるとの前提に立って操作化を進めるとしたら問題が含まれることになるであろう。たとえば、贈り物が手渡されることを前提として測定されたら日本における贈り物の相当部分は欠落してしまう。

そして、注意すべきは複数の国にまたがって共通の操作化がなされることが常に望ましいとはかぎらないことである。操作化を形の上で等しくすることだけに注意が払われる場合、それぞれの国での最も適切な操作化が達成できなくなる可能性がある。操作化にあたって重要なことは、いずれの国においてもその概念が十分に捉えられていること、つまり共通部分の捉え方において根本的な差異がないということであって、形式的に全く同一の操作化がなされていることではないのである。

概念の操作化はより具体的にはどのような項目が用いられるべきかということであり、項目相等性 (item equivalence) と呼ばれることも

ある (Hui & Triandis 1985)。消費者の態度や満足度のように複数の項目が用いられる場合、操作化における相等性は同じ項目が用いられるのかという点から検討されねばならない。

測定相等性 (measure equivalence)

概念の操作化における相等性の検討に関連する問題は測定そのものが相等するかということである。測定の相等性は測定に用いられる尺度の相等性 (scale equivalence)、言語の翻訳における相等性 (translation equivalence)、及び測定された数値そのものの相等性 (score equivalence) の三つの内容に分けて考えられる。もちろん、こうした各種の相等性については必ずしも研究者間で完全な合意がある訳ではない。各種相等性についての分類の仕方、用語の違い、同じ用語であっても意味範囲の相違と見解はまちまちである。本稿の記述は既存の文献に依拠しつつ筆者の考えに基づき整理したものであるが、概念的 성격のものから、より技術的性格なものへという階層順に従ったものである。

尺度相等性は複数の国で用いられる尺度が相等なものになっているかを問うことである。たとえば、ある国で態度の測定に間隔尺度が用いられているのに別の国では順序尺度になっていたり、「どちらでもない」といった中立的値が一方の国だけで用いられたりする場合には相等性が崩れることになる。

ただし尺度の相等性は操作化における相等性のところで検討された各項目について用いられる尺度が相等するか否かということであるから、それは選ばれる項目と密接に繋がったものである点に注意せねばならない。その意味で用いられる尺度は複数の国にまたがってできるだけ共通している方が望ましいとしても、それぞれの国において選ばれた項目との整合性を欠くことのないようにすることが大切である。

尺度の相等性は比較的研究者の目につきやすい形で確認が可能である。それに対して、より

微妙な形で実質的にこの尺度相等性を崩してしまうのは、異なる言語が用いられる場合の翻訳相等性である。たとえばSD法形式の5点尺度が用いられる場合、両極の形容詞の翻訳が不適切であるなら、本来は同じ尺度が意図されているにもかかわらず結果的には尺度は異なったものになってしまうことになる。この翻訳上の問題を避けるためには複数の言語に通じた研究者が細心の注意を払いつつ翻訳を行なうことと、第三者による逆翻訳 (back translation) による妥当性のチェックが不可欠である。翻訳の相等性の問題は確かに容易でない性格のものではあっても、こうした注意によってのり越えられない障害ではないと言われる (Sekaran 1983)。

尚、ここで測定相等性の中で論じられた翻訳相等性は項目相等性とも関連していることが明らかである。一つの国において特異な項目となってしまう事柄の場合、翻訳も容易でないということがあるからである。たとえば日本語の「甘え」という言葉は適切な英訳が困難であるといわれるが (土居 1971)、質問項目としても複数の国にまたがって行なうことに困難が伴うであろう。

測定相等性のうち最も相等性の確認の困難なものが測定された数値が同じものであるのかというスコア相等性である (Davidson & Thomson 1980)。スコア相等性は研究者によってはスカラー相等性 (scalar equivalence) と呼ばれることもある (Hui & Triandis 1985, Jaccard & Wan 1986)。スコア相等性の難しさは国別の測定値の平均や分散の差が測定せんとしている概念の量的な差に基づいているのか、それとも以上に述べた相等性の欠除が複合して生じているのかを識別できないだけでなく国別の回答の仕方の差のようなものが反映されている可能性があるからである。たとえばある商品の購買後の満足や不満の度合いを回答するにあたって、大げさな表現がなされる国と控え目な表現をする国とではスコアはそのまま比較できないことになる。

こうした点を考えると複数の国の平均値の差を単純に比較することは問題が多いことが明らかになる (Irvine & Carrol 1980)。もちろん平均値の比較は常にスコア相当性が欠如しているということではなく、とりあげられる平均値の性質によって相等性は異なっていると考えるべきである。消費者行動においても一般に行動そのものの比較は意識の比較よりもスコア相等性が高いと考えられる。意識の方が測定において文化差の影響が混入しやすいと考えられるからである。そして同じく何らかの意識についての平均値の比較と言っても、単一の項目の平均値より複数の項目の総合点の平均値の方がスコア相等性は高いと考えられる。幾つかの項目の総合によって極端なバイアスが相殺される可能性が考えられるからである。

では平均値の比較といった単純なスコアの相等性をおさえることが難かしいとすると、国際間の比較はどのようになされるべきであろうか。国際的な比較においてより高い比較可能性を有しているのは幾つかの変数間の関係を示すような数値である (Leung & Bond 1989)。たとえば、それぞれの国において回帰分析が行なわれ、回帰係数が比較される場合にはスコア相当性はかなり高いものと考えられる。それは複数の国で尺度相等性が成立しているかぎり、たとえばデータのスコア相等性がなくても回帰係数はほぼ同じ意味を持つと考えられる。複数の国での平均値の差は回帰式における定数項の差としてあらわれ、変数間の関係である回帰係数の差としてはあらわれないからである。又項目別の分散の大きさの効果を除去するには、変数を標準化して回帰係数を求めるというアプローチも考えられる。

データ収集の相等性 (equivalence in data collection)

以上の相等性は概念がどのように規定され、操作化され、そして測定されるかということに関しての相等性であったが、消費者行動の国際

どこまで共通しているのかを検討し得ると言う二点にあると言える。前者の有用性は消費者行動の国際比較においても構成概念の比較検討を含むのが普通であるから、比較されるそれぞれの国において構成概念が問題なく測定されているか否かの検討を可能にする点にある。後者はモデルが二つ以上の複数の国で共通した因子を持つのか、測定誤差には差がないのか、構造パラメータは同一とみなせるのかといった具合に段階的に統計的テストを行ない得ることである。

ここでは具体的なデータに基づいて、マルチ・サンプル分析によって先に論じた各種の相等性がどのように検討されるのかをみていくことにしよう。

ここでとりあげられるデータは日本と米国における学生のファーストフード・レストランに対する知識、態度、及び行動に関するものである^{注)}。比較に用いられるモデルは図-1に示されるように Fishbein と Ajzen (1975, 1980) の理由づけられた行動モデル (model of reasoned action) をベースとするものである。

ただし

B_i : ファーストフード・レストランで食事をすることが属性 (結果) i (たとえば便利な立地) をもたらす確かさ。信念と呼ばれる。

a_i : 属性 (結果) i の魅力度。

NB_j : 準拠集団 j がファーストフード・レストランで食事をすべきだと思っているとの考え。規範的信念と呼ばれる。

MC_j : 準拠集団 j の期待に従いたい動機づけ。

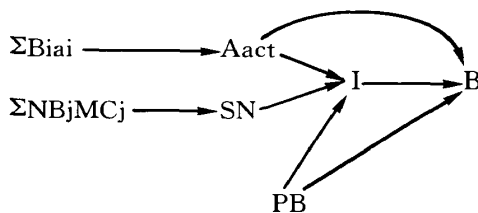


図-1 比較モデル

注) 尚、ここでのファーストフード・レストランでの食事は友人と一緒にの食事であり、 B_i と a_i を除く以下の変数はすべて友人と一緒にという文脈で測定されている。

Aact : ファーストフード・レストランで食事をすることに対する全体的評価。態度。

SN : 主観的規範、その学生にとって重要な人々 (諸準拠集団) がファーストフード・レストランで食事すべきだと思っているという信念。

PB : 最近の過去2週間におけるファーストフード・レストランでの食事回数。

I : 意図、向こう2週間にファーストフード・レストランで食事をする意図。

B : 行動、調査後2週間における実際のファーストフード・レストランでの食事回数。

図-1で Fishbein と Ajzen のモデルに追加的な形で用いられている変数は過去の行動 PB である。過去の行動はファーストフード・レストランの利用行動の中の主として習慣行動的な部分を把握する目的で加えられたもの (Bentler & Speckart 1979) である。そして、態度が意図をとおしてのみ行動に影響するものか否かを確認する意味で態度 A から行動 B への直接のパスも追加されている (Bagozzi 1981)。

日本と米国の学生のファーストフード・レストランの利用行動の比較はもちろん図-1のモデル全体について行なわれるわけであるが、ここでは簡単化のため主観的規範 SN から意図 I へのパスに着目することにしよう (Abe & Bagozzi 1993)。日米の学生のいずれについても意図は態度、主観的規範、過去の行動によって影響を受けることが考えられている訳であるが、集団主義的な文化的背景を考慮すると、日本の学生は主観的規範の役割が米国の学生にくらべて大きいことが予想されるからである。そこでここではこの SN から I へのパスについて日本と米国の同異を確かめるかたちでマルチ・サンプル分析を行なうことにする。

まず図-1に含まれる構成概念について日本と米国との間には概念的相等性が認められると思われる。ファーストフード・レストランという両国で極めて類似したサービス業の利用に関して図-1に含まれる構成概念は同じ意味内容を有していると考えてさしつかえないからであ

る。又、われわれは図-1のモデルに特定化される構造が日米両国にまたがって仮説としてあてはまることを考えている訳であるから機能的相等性も確認することができる。もちろん図-1のパスについて結果として推定されるパラメータが日米両国でかなり違ったものとなることは当然予想されるが、そうした差異を見出すことは比較研究の目的にこそなるものであっても、比較自体を無意味にするものではないから、機能的相等性に相反するものではない。あるいは一方の国において、あるパラメータが有意であって、他方の国ではそれが有意でないという結果が得られることも十分予想されるが、それも機能的相等性を崩すものではないと思わ

れる。モデルのテストにあたってははじめから違った特定化がなされなければならないような差異の場合に、機能的相等性が欠けると考えられるべきであろう。

構成概念の操作化における相等性は、経験的研究にあたって同一の測定項目を用いることができることで確認できる。具体的な測定項目が日本と米国で等しいもので差しつかえないということの一つの証しとして、信念Biについて行なった属性の抽出調査があげられよう。今回のデータ収集にあたって調査票の作成前に日本と米国でファーストフード・レストランに関する属性を抽出するための探索的調査を行なった。表-1は自由回答形式でファーストフード・レ

表-1 日米両国における属性の抽出

日本 38人		米国 41人	
利点			
サービスの早さ	23	サービスの早さ	21
使い易さ	20	便利な立地	18
座って話しができる	12	手頃な値段	17
便利さ	11	食べ物の品質	12
食べ物の品質	6	栄養	6
遅くまでやっている	4	清潔	6
軽食として丁度良い量	4	メニューの豊富さ	5
便利な立地	3	便利さ	4
メニューの豊富さ	3	割安	3
割安	2	雰囲気の良い	3
清潔	2	その他	3
雰囲気の良い	2		
従業員の態度	1		
その他	2		
問題点			
食べ物の質	27	食べ物の質	19
割高	22	割高	14
食事として量不足	11	脂肪/コレステロール	11
メニューの少なさ	9	混雑 (行列)	9
混雑	6	混雑 (座席なし)	7
包装の無駄	3	不潔さ	7
雰囲気	2	従業員の態度	6
従業員の態度	2	栄養のなさ	5
不潔さ	2	食事として量不足	5
立地	1	メニューの少なさ	4
その他	8	雰囲気	4
		包装の無駄	4
		立地 (駐車場のなさ)	3
		その他	3

レストランの利用にあたって利点と問題点と思われる項目(属性)を挙げてもらった結果である。

表-1から利点の順位について日米で多少の差異があるもののほぼ同じ項目があげられていることがわかる。本調査ではこの結果と過去の研究(Brinberg & Durand 1983)に基づき12の属性を用いることにした。それらは、早い給仕、食べ物の味の良さ、混雑、メニューの豊富さ、便利な立地、栄養、低脂肪とコレステロール、清潔、割安、手間のかからない便利さ、雰囲気、包装である。これらはそれぞれの魅力 a_i でウェイト付けされ、態度を規定する要因となる。

次に規範的信念 NB_j 及びそれに従いたい動機づけ MC_j にかかわる準拠集団として、友人と両親の二つの準拠集団が同様に探索的調査によって抽出された。

図-1のモデルに含まれるその他の構成概念については特別な探索的調査を必要と考えられなかったため、先験的に日米で共通の指標を決

めて用いることとした。態度は快適—不快、魅力的—非魅力的、有害—有益、の3項目、意図はありそうにない—ありそうだと確率の2項目が用いられる。過去の行動及び行動はそれぞれ2週間内にファーストフード・レストランを利用した回数である。これらの概念について日本と米国における操作化は相等性を想定しても問題ないと思われる。

図-1のモデルはLISRELモデルとして表示すると図-2のようになる。図-2で丸で囲まれた変数は図-1における構成概念であるが、四角で囲まれた変数は測定のための指標である。日米で同一のLISRELモデルが用いられることは、同一の操作化がなされ、項目相等性が成り立つことを意味している。

尚、図-2で構成概念と指標との矢印 λ は因子負荷量で両者の結びつきを示す。単一の指標が用いられる場合にはその指標で構成概念が捉えられると考えられているため、因子負荷量は1.0である。複数の指標が用いられる場合、そ

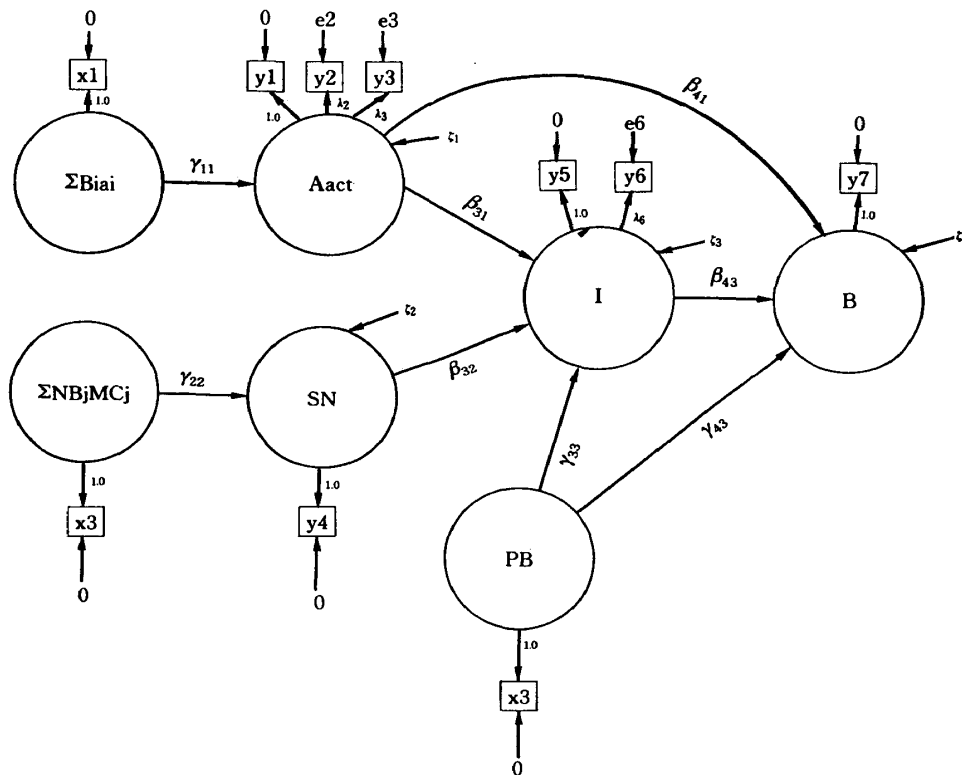


図-2 LISRELモデル

のうちの一つは概念と指標の単位を同一とするという意味で1.0に定めることができる。

図-2で構成概念間のパス β (内生変数間)と γ (外生変数と内生変数との間)は先にとりあげた機能的相等性にあたる。これらのパスを推定し国際間の差異を探っていくことが比較研究の目的である。

図-2で指標へのもう一つの矢印 e はランダムな性質を持つと考えられる測定誤差である。指標が単独の場合には測定誤差を0と想定せざるを得ないが、複数の場合には測定誤差を押えての分析が可能となる。それは構成概念の測定におけるノイズをとることであり、構造パラメータ β, γ のより明確な把握が可能となることを意味している。

測定誤差 e の分散の大きさと因子負荷量 λ についての比較検討は測定の相等性を内容的に検討することにあたっている。本研究ではすべての項目について日本と米国で同一の尺度が用いられている。同一の尺度を用いることに特別な問題が発生することが予想されないから尺度相等性は高いことが想定されている。それでも、もしLISRELモデルによって推定された因子負荷量や測定誤差の分散が似かよった値になるならば、それだけ測定が相等な尺度に基づいてなされているという確信は強くなるであろう。因子負荷量が同じであるということは因子構造が相等しいということであり、概念の相等性を支持する意味もあわせ持っている。

測定誤差の分散が等しいことは用いられた尺度が同程度に信頼できることを意味している。比較される国の中で尺度の信頼性が大きく異なることは望ましいことではない (Davis, Douglas & Silk, 1981) から、測定誤差の分散の大きさも尺度の相等性にとって検討されるべきポイントであることがわかる。

以上で明らかなように因子負荷量と測定誤差についてのこうした検討は複数の指標が用いられたときにのみ可能であって、単一の指標しか持たない構成概念についてはできるものではな

い。その意味で複数の指標を用いた分析を可能とするLISRELが国際比較においても有用であることがわかる。

尚、マルチ・サンプル分析に直接かかわるものではないが、今回の分析において測定相等性のもう一つの側面である翻訳相等性は英語で作成した調査票を日本語に翻訳し、それを第三者によって英語に再翻訳し、原文との同異を確かめ翻訳の妥当性を高める方法をとることにした。サンプルとしては日本では専修大学、青山学院大学、横浜国立大学のそれぞれ経営学部で収集された419名のうち、不完全な回答を除いた288名を用いた。米国ではミシガン大学の文理学部 of 学生246名のうち有効となった179名である。いずれも授業に出席した学生を対象とした便宜標本であり、学部も異なっているから厳密なサンプル相等性には問題が含まれている。ただし、こうした不完全性によって両サンプルの間でファーストフード・レストランの利用に関して決定的な攪乱要因が混入してくるとは考えられないから今回の比較目的には十分使えるものであると思われる。

データ収集方法については、いずれも授業に出席した学生に調査票を配布、その場で記入されたものを回収するという方法をとっており、方法的に全く同一である。

そこでこうした相等性の確認をふまえて、今回の分布の関心事である主観的規範SNから意図Iへのパス β_{32} を中心に日米のサンプルで差があるか否かをみてみることにしよう。既にスコア相等性のところでみた如く、日米のサンプル間で測定されたスコアの平均値は相等性が低いにしても、変数間の関係として計算されたパラメータの間には相等性が高いと考えられる。二つのサンプルの間の類似性を確めるために以下のような技術的な仮説を設定することにしよう。

(仮説1) H_{null} : 二国間でパラメータは全く異なる。

(仮説2) $H_{\beta_{32}}$: 二国間で β_{32} だけ共通。

(仮説3) H_A : 二国間で因子負荷量が共通.

(仮説4) $H_{A\beta_{32}}$: 二国間で因子負荷量と β_{32} が共通.

(仮説5) $H_{A\theta}$: 二国間で因子負荷量と測定誤差の分散が共通.

(仮説6) $H_{A\theta\beta_{32}}$: 二国間で因子負荷量, 測定誤差の分散, 及び β_{32} が共通.

(仮説7) $H_{A\theta\beta_{32}}$: 二国間で β_{32} を除くすべてのパラメータが共通.

(仮説8) $H_{A\theta\beta_{32}}$: 二国間ですべてのパラメータが共通.

(仮説9) H_{Σ} : 二国間でパラメータ推定に用いられる分散, 共分散行列 Σ が同一.

マルチ・サンプル分析を用いると日米の両サンプルの間でどの仮説が採択もしくは棄却されるかをテストすることができる. ここにあげた仮説は仮説1から9へと進むに従ってゆるやかな仮説から厳しい仮説へと移ることになるから, どこまで日米間で共通しているかを探ることが可能である. たとえば仮説7は採択されるが仮説8は棄却されるという場合, β_{32} を共通とすか否かだけでモデルの採択が決まってくるわけであり, β_{32} を除く他のパラメータはすべて共通であっても β_{32} は共通ではないと結論づけられることになる. そしてもし, 仮説の採棄が仮説7と仮説5との間で分れるとするならば, 両サンプルは構造パラメータ β と γ を共通としていないことが明らかとなる.

もちろん, ここでとりあげた仮説の段階は大幅なものであって, 推定されるパラメータの一つ一つの同異を確かめるという小刻みなアプローチをとることも可能である. そして仮説1が棄却される場合にはマルチ・サンプル分析が適用できないことを意味する. 仮説1は日本と米国でそれぞれ独自のパラメータを求めるというものであるから, それが採択とならないことは, 日米いずれかあるいは両方のモデルのあてはまりが低いことになるからである. すなわち, マルチ・サンプル分析は比較される複数の国でモデルが採択されるときにのみ意味を持ってく

ることがわかる. 仮説9はモデルの特定化を行なう前の投入データの同異をテストするもので最も厳しいテストである. モデルの採否にはモデルの適合度を示す χ^2 値の得られる確率 P が用いられる. 一般に $P < 0.1$ のときはモデルを棄却, $P \geq 0.1$ のときはモデルを採択するのが普通である (Bagozzi, 1980).

ファーストフード・レストランを対象とした今回の研究では日本と米国のサンプルについてそれぞれ LISREL による分析を行なったところ, 日本のサンプルは $\chi^2 = 46.02$ (自由度29, $n = 288$) $P = 0.023$, 米国のサンプルは $\chi^2 = 96.42$ (自由度29, $n = 179$), $P = 0.000$ でいずれも棄却されてしまうからマルチ・サンプル分析を行なうことはできない. そこで日米の両サンプルについて心理的尺度の一つである INDCOL (個人主義—集団主義) (Hui, 1984) 尺度のうち代表的とされる8つの項目 (Triands, 1988) を用いてサンプルを二分し, 平均スコア以下のサブサンプル間での比較を行なうことにする. その場合, 日本のサブサンプルは $\chi^2 = 24.97$ (自由度29, $n = 167$), $P = 0.680$, 米国のサブサンプルは $\chi^2 = 38.40$ (自由度29, $n = 91$), $P = 0.114$ となり, いずれもモデルは採択となるからである. すなわち, 日米の両サンプルとも集団主義的傾向の低い (個人主義的) サブサンプルについてはモデルのあてはまりが良いことになる. ここで集団主義の8項目の平均値は日本が27.08で米国が26.25である (5%で有意差あり), この数値からするかぎり同じく平均以下のサブサンプルであるといっても日本側の方が僅かにより集団主義的であることになる. ただし, この平均値の差は8項目の総合点であるところからある程度の比較可能性を持つことは期待されるものの, 厳密な意味でのスコア相等性を有していると考えられない. したがって, 統計的有意差だけをもってただちに日本のサンプルの方が集団主義的と断定することは控えられるべきである.

日米両サンプルについて, それぞれ推定され

表-2 パラメータ推定の結果

	米国	n=91 標準化	日本	n=167 標準化
λ_{y1}	1.0	.837	1.0	.763
λ_{y2}	1.298***	1.087	1.152***	.878
λ_{y3}	.784***	.656	.850***	.648
λ_{y4}	1.0	1.459	1.0	1.502
λ_{y5}	1.0	1.576	1.0	1.275
λ_{y6}	1.672***	2.635	1.470***	1.877
λ_{y7}	1.0	1.393	1.0	1.472
λ_{x1}	1.0	64.328	1.0	67.092
λ_{x2}	1.0	17.758	1.0	12.279
λ_{x3}	1.0	1.824	1.0	2.869
β_{31}	.319*	.170	.431***	.258
β_{32}	.163*	.151	.051	.060
β_{41}	.147	.089	-.368**	-0.191
β_{43}	.236**	.267	.512***	.443
γ_{11}	.005***	.346	.002*	.140
γ_{22}	.016*	.194	.001	.011
γ_{33}	.490***	.567	.139***	.313
γ_{43}	.343***	.449	.071**	.138
	信頼性係数	var (e)	信頼性係数	var (e)
y_1	.514	.662***	.627	.345***
y_2	.928	.091	.774	.226***
y_3	.317	.927***	.474	.466***
y_4	1.0	0	1.0	0
y_5	.720	.968***	.600	1.082***
y_6	.879	.958	.661	1.803***
y_7	1.0	0	1.0	0
χ^2		38.40		24.97
d. f.		29		29
p		.114		.680
GFI		.932		.971

* 10%で有意
 ** 5%で有意
 *** 1%で有意

るパラメータを表示すれば表-2のとおりである。

表-2から関心事とするパス β_{32} についてみると米国のサンプルでは β_{32} が0.163 (10%で有意), 日本のサンプルでは β_{32} が0.051 (10%レベルでも有意でない) となっていることがわかる。この結果からするかぎり, 仮説とは逆に米国の学生の方が意図の形成にあたって主観的規範が有意な関係 (10%レベルであるが) を持っていることがわかる。

その他のパラメータについて因子負荷量が両サンプルで比較的似かよった値となっていることがわかる。それ以外のものには差がみられるから, マルチ・サンプル分析のための仮説では仮説3までは採択となり, 仮説5以上は棄却となることが予想される。

LISRELによるマルチ・サンプル分析では複数のサンプルについて同時にパラメータを推定する方式がとられる。その際パラメータの共通性の条件を各仮説に応じて制約条件として与

表-3 マルチ・サンプル分析の結果

仮説	χ^2	自由度	P	決定
(仮説1) H_{null}	63.37	58	0.293	採択
(仮説2) $H_{\beta_{32}}$	63.56	59	0.319	〃
(仮説3) H_{Λ}	64.69	61	0.349	〃
(仮説4) $H_{\Lambda\beta_{32}}$	64.76	62	0.380	〃
(仮説5) $H_{\Lambda\theta}$	83.67	66	0.070	棄却
(仮説6) $H_{\Lambda\theta\beta_{32}}$	83.70	67	0.082	〃
(仮説7) $H_{\Lambda\theta\beta_{32}\Gamma}$ (β_{32})	118.94	73	0.001	〃
(仮説8) $H_{\Lambda\theta\beta_{32}\Gamma}$	119.60	74	0.001	〃
(仮説9) H_{Σ}	123.68	55	0.000	〃

えればよいだけである。そしてモデルの適合度は複数のサンプル全体として得られることになる。いま、各仮説に応じたマルチ・サンプル分析の結果を適合度について表示すれば表-3のようになる。

表-3から仮説4までが採択となることがわかる。すなわち両サンプルについて因子負荷量と β_{32} を等しいとするモデルまでは採択されるわけである。二つのサンプルの間で β_{32} が等しいという条件が設定されることはモデルの適合度を下げることにはならないこともわかる。実際のところ仮説1と仮説2、仮説3と仮説4、仮説5と仮説6との対比から β_{32} の追加によってP値は増加しているのがわかる。 β_{32} の追加による χ^2 値の名目的な増加よりも自由度が増加する効果の方が大きいからである。したがって、低い集団主義のサブサンプルについては日米の間で β_{32} を同一と考えてもさしつかえないことになる。そしてこの結論は両サンプルを別々に分析した表-2の結果を補完し深めることになると思われる。表-2では β_{32} について推定値の僅かな差が観察されたものの、その差をテストするものではなかったからである。

表-3から、二つのサブサンプルは測定誤差の分散と β_{32} 以外のものを一括した構造パラメータにおいて差のあることがわかる。このうち測定誤差の分散が違うことは先述のように測定の相等性を割引くものと考えられる。ただし、測定は日米いずれかで一方的に旨くいつているということではない。信頼性係数は態度につい

ては日本のサンプルが大であり、意図については米国のサンプルが大となっているからである。構造パラメータ β と γ における差異は今回の測定が問題なく行なわれているとした場合日米間での消費者行動の差異をあらわしているものになると考えられる。その意味で構造パラメータの比較は概念相等性、操作化における相等性、測定相等性、データ収集における相等性を前提としてはじめてなされるものと考えられる。しかしここでも先に論じたように推定された構造パラメータが大幅に異なる場合には、比較される概念そのものの役割が異なること、すなわち機能的相等性が成り立っていないことを示唆するものとなる。たとえば今回の研究で用いられた態度は因子構造はほぼ等しく概念的相等性を有していると考えられるものの、行動へのパス β_{41} に関しては大幅な差異（日本の場合、負）となっているから、やや異なった役割を持ったものとして捉えられることになる。したがって比較可能性の検討は概念的相等性→操作化における相等性→尺度の相等性→スコアの相等性という、概念的なものをベースとしてより技術的なものが検討されるという階層構造が考えられている（Hui & Triadis, 1985）わけであるが、それは研究の進められる原則の流れであって、実際には結果として具体的に得られた数値からより概念的レベルの相等性に関する疑問が出されるというフィードバックがなされるものと考えられよう。この問題は構成概念の測定が妥当であるかという検討が一般に構成概念そのもの

は妥当であるとの想定のもとになされることに似ていると言えよう (阿部, 1987)。

その場合でも測定結果が予測されたものと大きくずれてしまうことは測定の問題としてだけでなく、構成概念そのものの問題である可能性を有しているのである。

今回の研究では因子構造の検討のからみにおいて日米両サンプルの態度の果たす役割上の差異が検討されるべきことになろう。特に今回の場合、日本のサブサンプルで態度から行動への直接のパス β_{41} がマイナスで有意となっていることは注目し得る。態度は意図をとおして行動にプラスに影響しているが、それでも全体としての影響はマイナスの方が勝るから、ファーストフード・レストランに対する態度がよい程、その利用回数が少ないという結果となっているからである。集団主義傾向の低い日本のサブサンプルについてなぜこうした結果が得られたのかは別途の検討が必要である。

おわりに

以上、比較可能性を概念的相等性、操作化における相等性、測定相等性、データ収集における相等性の四つの側面に分けて考察を行ない、つづいて LISREL におけるマルチ・サンプル分析によってどのようにこうした相等性が確認されるのかをみてきた。消費者行動の交叉文化的研究におけるマルチ・サンプル分析の有用性は、それが構成概念妥当性の検討を可能にすることと、複数のサンプルがどこまで類似しているのかを押えることができる点にある事が明らかとなった。このうち後者の有用性は、複数のサンプルにまたがって共通の概念、共通の項目及び共通の尺度が用いられているとき、すなわちスコア相等性を除く測定相等性が成立しているときにのみ認められるものであった。又、比較のなされるサンプル毎のモデルの適合度の高いことも必要条件であることが明らかとなった。そしてそのことはマルチ・サンプル分析が相当程度に相等性の高いサンプル間での比較の場合

に限って有用なアプローチであるという限界を持つことを意味すると同時に、逆にその適用がそうした相等性についての検討を必然ならしめるが故により厳密な比較可能性の検討を促すという側面も有しているといえよう。今後、こうしたフォーマルなかたちでの交叉文化的消費者研究がより多くなされていくことが望まれるのである。

付記 本研究にあたりデータ収集を可能として下さったミシガン大学の R. P. Bagozzi 教授、専修大学の奥田和彦教授、青山学院大学の田中正郎助教授にお礼を申し上げます。又、本稿は吉田秀雄記念事業財団の平成三年度の研究助成を受けてなされた研究の一部であります。ここに心より感謝申し上げます。

参考文献

- 1) 阿部周造 (1984) 「消費者行動の国際比較——その予備的考察——」 横浜経営研究, 4 巻, 115-122.
- 2) 阿部周造 (1987) 「構成概念妥当性と LISREL」 奥田和彦, 阿部周造編『マーケティング理論と測定』 中央経済社 27-46.
- 3) Abe, Shuzo, and Richard P. Bagozzi (1993), "International Comparison of Consumer Behavior," マーケティング・サイエンス, 第 1 巻, 収録予定.
- 4) Bagozzi, Richard P. (1980), *Causal Models in Marketing*, John Wiley & Sons.
- 5) Bagozzi, Richard P. (1981), "Attitudes, Intentions, and Behavior: A Test of Some Key Hypotheses," *Journal of Personality and Social Psychology*, Vol. 41, 607-627.
- 6) Bentler, P. M., and G. Speckart (1979), "Models of Attitude Behavior Relations," *Psychological Review*, Vol. 86, (5), 452-464.
- 7) Berent, P. H. (1975), "International Research is Different," in E.M. Mazze (ed.), *Combined Proceedings, American Marketing Association*, 293-297.
- 8) Briberg, David, and Jody Durand (1983), "Eating at Fast-food Restaurants; An Analysis Using Two Behavioral Intention Models," *Journal of Applied Social Psychology*, 10, 459-472.
- 9) Dant, Rajiv P. and James H. Barnes, Jr. (1988),

- “Methodological Concerns in Cross-Cultural Research: Implications for Economic Development,” in Erdogan Kumuc and Fuat Firat (eds.), *Research in Marketing*, JAI Press Inc, 149–171.
- 10) Davidson, Andrew R., and Elizabete Thomson (1980), “Cross-Cultural Studies of Attitudes, and Beliefs,” in Harry C. Triandis and Richard W. Brislin (eds.), *Handbook of Cross-Cultural Psychology: Social Psychology*, Vol. 5, Allyn & Bacon Inc., 25–71.
 - 11) Davis, H.L., S.P. Douglas, and A.J. Silk (1981), “Measure Unreliability: A Hidden Threat to Cross-National Marketing Research?” *Journal of Marketing*, Vol. 45, No. 2, 98–109.
 - 12) 土居健郎 (1971) 『甘えの構造』弘文堂.
 - 13) Fishbein, Martin, and Icek Ajzen (1974), *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Addison-Wesely.
 - 14) Fishbein, Martin, and Icek Ajzen (1980), *Understanding Attitudes and Predicting Social Behavior*, Prentice-Hall.
 - 15) Hui, c. H. (1984), “Individualism-Collectivism: Theory, Measurement and Its Relation to Reward Allocation,” Un-published Doctoral Dissertation, The University of Illinois.
 - 16) Hui, C. Harry, and Harry C. Triandis (1985), “Measurement in Cross-Cultural Psychology: A Review and Comparison of Strategies,” *Journal of Cross-Cultural Psychology*, Vol. 6, No. 2, 131–152.
 - 17) Irvine, Sid H., and William K. Carroll (1980), “Testing and Assessment across Cultures: Issues in Methodology and Theory,” in Harry C. Triandis and John W. Berry (eds.), *Handbook of Cross-Cultural Psychology, Methodology*, Vol. 2, Allyn and Bacon, 181–244.
 - 18) Jaccard, James and Choi K. Wan (1986), “Cross-Cultural Methods for The Study of Behavioral Decision Making,” *Journal of Cross-Cultural Psychology*, Vol. 17, No. 2, (June), 123–149.
 - 19) Jöreskog, Karl G., and Dag Sörbom (1989), *LISREL 7: A Guide to the Program and Applications*, 2nd Edition, SPSS Inc..
 - 20) Leung Kwok, and Michael Harris Bond (1989), “On the Empirical Identification of Dimensions for Cross-Cultural Comparisons,” *Journal of Cross-Cultural Psychology*, Vol. 20, No. 2, 133–150.
 - 21) Mayer, C.S. (1980), “Multinational Marketing Research: Methodological Problems,” in H. Thorelli, and H. Becker (eds.), *International Marketing Strategy: Revised Edition*, Pergamon Press, 162–172.
 - 22) Sekaran, Uma (1983), “Methodological and Theoretical Issues and Advancements in Cross-Cultural Research,” *Journal of International Business Studies*, Fall, 61–73.
 - 23) Triandis, Harry C. (1988), “Collectivism v. Individualism: A Reconceptualization of a Basic Concept in Cross-Cultural Social Psychology,” in C. Bagley and G. Verma (eds.), *Cross-Cultural Studies of Personality, Attitude and Cognition*, London, Macmillan.

[あべ しゅうぞう 横浜国立大学経営学部教授]