外国語教育メディア学会 (LET) 関西支部 メソドロジー研究部会 2011 年度報告論集
水本 篤 (pp. 63–73)

質問紙調査における相関係数の解釈について

水本 篤
関西大学

概要

本稿では，Dörnyei (2001) の「質問紙を用いた相関係数で，0.3–0.5 であれば，ある程度意味がある」という主張に対して，相関係数の希薄化の観点から，理論的な説明に加えて，R を用いたシミュレーションで経験的に検証を行った。シミュレーションの結果，信頼性係数が低い尺度を使って相関係数を求める場合には，相関係数の希薄化が起こり，相関係数で 0.3–0.5 の値が得られた場合には，ある程度の意味を持つ相関があると解釈できるということがわかった。

Keywords: 相関係数，質問紙，信頼性係数，相関係数の希薄化，シミュレーション

1. はじめに

相関分析をはじめとする相関係数を利用した分析は，研究論文では必ず目にするため，最も広く行われている分析であると言えるだろう。相関係数の大きさについての解釈の基準は，研究対象や，研究に用いた測定道具（項目数・変数の種類）や，何と何の相関係数を調べているのかなどという，文脈によって違ってくるものであるが，一般的には以下の表1のように判断することが多い。

表 1
相関係数の大きさと解釈（小塩, 2004, p. 29）

.00 〜 ± .20	ほとんど相関がない（.00 は無相関）
.20 〜 ± .40	低い（弱い）相関がある
.40 〜 ± .70	かなり（比較的強い）相関がある
.70 〜 ±1.00	高い（強い）相関がある

しかし, 動機づけなどの学習者要因を質問紙調査によって得られた相関係数を解釈する際には, Dörnyei (2001) では “In L2 motivation studies, the usual strength of the meaningful relationships detected is between 0.30 and 0.50” (p. 224) としており，表1の一般的な基準から考えると，質問紙調査における相関係数で意味があるとされる 0.3–0.5 という値は少し低い値だと感じるだろう。つまり，（質問紙であれ，テストであれ）ある2つの尺度から得られた相関係数が $r = .30$ の場合，研究者によっては表1の一般的な基準から「低い相関がある」と解釈し，別の研究者は「ある程度の相関がある」と解釈するような状況もあり得るということになる。

本稿では，このような現象を引き起こす原因である「相関係数の希薄化」について，理論的な説明を加え，Rを用いたシミュレーションで経験的に検証を行う。

2. 相関係数の希薄化

「動機づけ」や「英語力」は目に見えない概念であり，測定を行う者が「このような概念を測定しているであろう」と想定して測定しなければならない。そのため，質問紙でも，テストでも，測定したいもの（構成概念）を1つの項目で測定するということは不可能である。例えば，「一般動詞の使い方が理解できているか」ということを確認するために，1問の文法問題に対して，正解したか不正解したかだけでは，実力で正解したのか，問題の意図を読み違えて不正解だったのかわからないため，その能力を正しく測定することができない。そのため，問題数をある程度用意して，その正答数の合計値（もしくは平均値）で「一般動詞の使い方が理解できているか」ということを測定する必要がある。

質問紙でも同様に，「英語を学ぶ意欲がどの程度あるか」ということを測定するためには，1つの項目では，「まったく当てはまらない」を（1）とし，「とてもよく当てはまる」を（5）として選ぶような5件法で，3を選ぶのか4を選ぶのかというのは，そのときの気分や環境，その他の考えうるすべての影響による誤差が含まれてしまう。そのため，一般的には3項目以上の項目を用意し，その平均値（もしくは合計値）を求めることで，その構成概念を測定するという方法を用いる（Dörnyei, 2003）。

表1はそのような質問紙の下位尺度の例である。この下位尺度では，英語学習に対する外発的動機づけを測るものとする。実際の質問紙調査の際には，その他の内発的動機づけなどの下位尺度があり，それらすべてを含めて，尺度（質問紙）として使用される。表1の下位尺度には，質問項目が6つあり，5件法で回答例のような回答パターンが得られたとき，$(3 + 2 + 3 + 4 + 3 + 4) / 6 = 3.1666...$ で尺度値が，3.167 という計算になる。

このような質問紙の下位尺度と，別の下位尺度や，テストの得点との相関係数を求める場合に，注意しなければならないのが，「相関係数の希薄化」（attenuation of correlation coefficients）である。相関係数の希薄化は，相関係数を計算する際に，尺度ごとの信頼性係数が1ではならないことから起こる現象である。測定において，信頼性係数が1になることはないので，誤差を含んだ値を用いて，相関係数を求めるために，関係が薄まってしまうのである（田中・前田, 2004）。

表 2

質問紙における下位尺度（外発的動機づけ）の例

質問項目 私が英語の勉強をするのは...	回答例	尺度値
1. よい成績を取ることが自分にとって重要であるからである	3	
2. 他の人よりもよい成績を取りたいと思うからである	2	
3. 社会的に認められる人物になるためである	3	3.167
4. 資格試験などでよりよい成績を収めたいからである	4	
5. 英語を勉強することが決まりのようなものだからである	3	
6. 将来よりよい仕事に就くためである	4	

相関係数の希薄化を修正して，それぞれの尺度の信頼性係数が 1 であった（誤差をもたない）場合に得られる信頼性係数の推定値を求めるために，以下のような希薄化の修正公式を用いることができる。

$$\text{修正された相関係数} = \frac{\text{観測された（修正なしの）相関係数}}{\sqrt{\text{尺度 A の信頼性係数} \times \text{尺度 B の信頼性係数}}}$$

この式からわかるように，例えば，尺度Aと尺度Bのそれぞれにおいて，信頼性係数（クロンバックのアルファ）が0.6程度で，観測された相関係数が0.3であった場合は，修正された相関係数は0.5となり，実際の相関（母相関）の6割程度の値になってしまう。

信頼性係数を高める要因の一つとして，問題数が多いことが前提であるため，質問紙の下位尺度では項目数が3項目というように少なく，信頼性係数も0.6−0.7程度の尺度もあるため，Dörnyei (2001) の主張するように， 0.3–0.5の値が得られた場合には，ある程度の意味を持つ相関があると解釈すると考えられる（表3参照）。ただし，相関係数の希薄化の修正は，修正後の値が1を越えることもあり，論文などで報告されることは少ない。

相関係数の希薄化に対する最善の解決策としては，構造方程式モデリング（structural equation modeling: SEM）の検証的因子分析（confirmatory factor analysis: CFA）を用いることである。SEMでは，「誤差をコントロールして希薄化されない正しい推定値を求めることができる」（狩野, 2002, p. 147）とされており，問題なく修正ができるため，サンプルサイズが（200以上が望ましいが）100以上である場合には，積極的にSEMを使うべきである。

表 3

信頼性係数と実際の相関，観測される相関係数

2 つの尺度の信頼性係数が 0.6 の場合		2 つの尺度の信頼性係数が 0.7 の場合	
実際の相関（母相関）	観測される相関係数	実際の相関（母相関）	観測される相関係数
.100	.036	.100	.049
.200	.072	.200	.098
.300	.108	.300	.147
.400	.144	.400	.196
.500	.180	.500	.245
.600	.216	.600	.294
.700	.252	.700	.343
.800	.288	.800	.392
.900	.324	.900	.441

次のセクションでは，観測された相関係数，希薄化修正された相関係数，SEMを用いた場合の相関係数の関係を，Rを使ったシミュレーションで検証する。分析にはすべて R version 2.14.2を用いた。

3. シミュレーションによる検証

3.1 発生させたデータ

項目 1〜5 の 5 項目が 1 つの下位尺度（因子），項目 6~10 の 5 項目がもう 1 つの下位尺度（因子）となるように R を使って質問紙の回答データを正規分布に従う乱数で発生させた。その際，それぞれの下位尺度の信頼性係数が 0.65 程度で 2 因子構造になるように，青木繁伸先生のサイトのプログラム（http://aoki2.si.gunma-u.ac.jp/R/gendat.html）を用い，表 4 のような項目間の相関を持たせた。

表 4

シミュレーションで発生させた項目間に持たせた相関

	項目 1	項目 2	項目 3	項目 4	項目 5	項目 6	項目 7	項目 8	項目 9	項目 10
項目 1	–									
項目 2	0.3	–								
項目 3	0.3	0.3	–							
項目 4	0.3	0.3	0.3	–						
項目 5	0.3	0.3	0.3	0.3	–					
項目 6	0.2	0.2	0.2	0.2	0.2	–				
項目 7	0.2	0.2	0.2	0.2	0.2	0.3	–			
項目 8	0.2	0.2	0.2	0.2	0.2	0.3	0.3	–		
項目 9	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	–	
項目 10	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3	–

上記までの手順で発生させたデータは平均値が 0，標準偏差が 1 になるように標準化されているため，5 件法の質問紙の回答と同じようになるように，Rcommander に含まれている数値変数を区間に分ける関数で 1, 2, 3, 4, 5 の 5 段階に変換した。データは乱数で発生させているので，毎回，数値は若干変動するが，例として，サンプルサイズを 200 にして発生させたデータの記述統計は表 5 の通りである。このデータについて，尺度値を用いた相関係数，希薄化修正された相関係数，図 1 のような SEM（CFA）を用いた場合の相関係数をまとめたものが表 6 になる。データを正規分布に従って発生させているため，希薄化修正された相関係数と SEM による相関係数がほとんど同じであることがわかる。

表 5

発生させたデータの記述統計

	項目数	*n*	平均	標準偏差	最小	最大	歪度	尖度	クロンバック α
尺度 A	5	200	3.010	0.589	1.600	4.400	-0.070	-0.259	.669
尺度 B	5	200	2.920	0.562	1.400	4.800	0.075	0.255	.636

Note. 乱数で発生させているため，数値は毎回変動する。

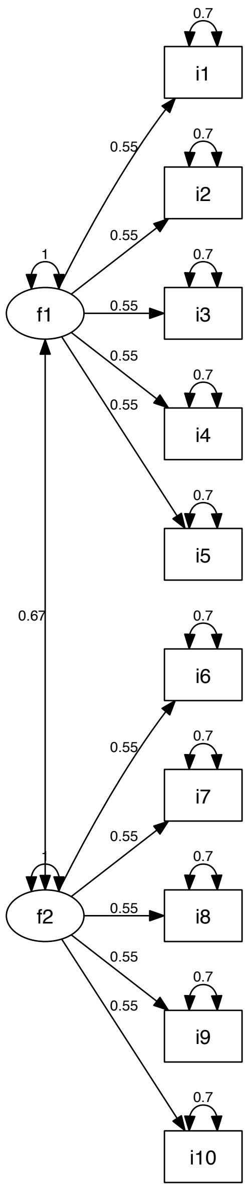

図 1.　検証的因子分析

表 6

データ（表 5）から得られる相関係数

	相関係数	希薄化修正された相関係数	SEM による相関係数
尺度 A 尺度 B	.434	.665	.667

3.2　シミュレーション

前項のように乱数で発生させたデータは，毎回変動するため，同じ処理を何百回，何千回と繰り返し，その平均を求めることで，どのような傾向があるかを調べるモンテカルロ・シミュレーションを実施した（シミュレーションの詳細については，山田・杉澤・村井, 2008を参照）。

今回のシミュレーションでは，相関係数が0.5のときに，検定力が0.8以上となる最低サンプルサイズである30から，60, 120, 240, 480とサンプルサイズを倍に変化させ，それぞれで10,000回データを発生させて，それぞれの尺度の信頼性係数，尺度値を用いた相関係数，希薄化修正された相関係数の平均を求めた（SEMによる相関係数は変化しない）。その結果を表7と図2に示す（使用したRのコードはAppendix を参照）。正規分布に従ったデータで，10,000回の繰り返しを行った平均であるため，数理的な説明と同じく，希薄化修正された相関係数とSEMによる相関係数がほとんど同じになっている。

表 7

モンテカルロ・シミュレーションの結果（10,000 回の平均）

n	尺度 A（5 項目）の信頼性係数	尺度 B（5 項目）の信頼性係数	尺度値による相関係数	希薄化修正された相関係数	SEM による相関係数
30	.657	.656	.440	.671	.667
60	.654	.653	.437	.670	.667
120	.651	.651	.435	.669	.667
240	.647	.647	.432	.668	.667
480	.643	.643	.430	.668	.667

Note. *n* は発生させたデータのサンプルサイズ。希薄化修正された相関係数は，表中の尺度 A，尺度 B，尺度値による相関係数を用いて計算したものではなく，シミュレーション後の平均値を掲載している。

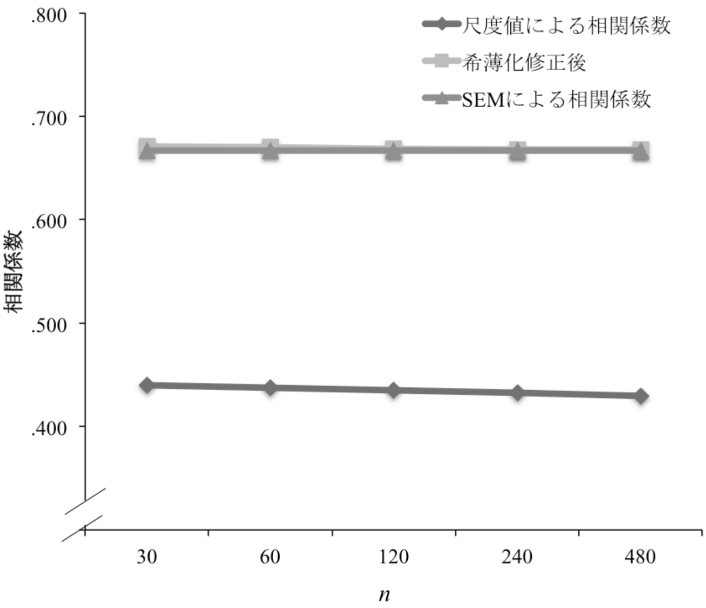

図 2.　モンテカルロ・シミュレーションの結果（10,000 回の平均）

4.　まとめ

本稿では，Dörnyei (2001) の「質問紙を用いた相関係数で，0.3–0.5であれば，ある程度意味がある」という主張に対して，相関係数の希薄化の観点から，理論的な説明に加えて，Rを用いたシミュレーションで経験的に検証を行った。

シミュレーションの結果，理論的な説明と同じく，信頼性係数が0.6–0.7程度の低い値の尺度を使って相関係数を求める場合には，誤差によって相関係数の希薄化が起こり，相関係数で 0.3–0.5 の値が得られた場合には，ある程度の意味を持つ相関があると解釈できるということがわかった。これらの結果から，相関係数を使った分析と解釈を質問紙調査で行う際には，以下の点に注意すべきであると言えるだろう。

- 信頼性係数を高めるために，下位尺度に含まれる項目数をできるだけ多くする。
- 信頼性係数が低い尺度を用いて相関係数を求める場合，希薄化の修正も参考として確認する。
- 質問紙による尺度を使う場合（テストでももちろん同様に），論文中に信頼性係数を明記する。
- （本稿で用いたデータは，正規分布に従って人工発生させたものであるため，希薄化の修正もSEMの結果とほとんど同じ値が得られたが，実際のデータではそのようにならないため）SEMが使える場合には積極的に使用する。

参考文献

Dörnyei, Z. (2001). *Teaching and researching motivation*. Harlow: Pearson Education Limited.

Dörnyei, Z. (2003). *Questionnaires in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

田中博晃・前田啓朗 (2004). 「外国語学習動機研究における構成概念 'amotivation'：測定の妥当性検証とネガティブな質問項目の影響」*JLTA (Japan Language Testing Association) Journal, 6*, 128–139.

山田剛史・杉澤武俊・村井潤一郎 (2008). 『R によるやさしい統計学』 オーム社.

Appendix 使用した **R** コード

```
# 10,000 回のシミュレーション
library(Rcmdr)   # Rmdr パッケージ使用
library(ltm)        # ltm パッケージ使用

N <- 10000
alphax <- numeric(length=N)
alphay <- numeric(length=N)
correl <- numeric(length=N)
adjust <- numeric(length=N)
for(i in 1:N){
r <- matrix(c(
1, 0.3, 0.3, 0.3, 0.3, 0.2, 0.2, 0.2, 0.2, 0.2,
0.3, 1, 0.3, 0.3, 0.3, 0.2, 0.2, 0.2, 0.2, 0.2,
0.3, 0.3, 1, 0.3, 0.3, 0.2, 0.2, 0.2, 0.2, 0.2,
0.3, 0.3, 0.3, 1, 0.3, 0.2, 0.2, 0.2, 0.2, 0.2,
0.3, 0.3, 0.3, 0.3, 1, 0.2, 0.2, 0.2, 0.2, 0.2,
0.2, 0.2, 0.2, 0.2, 0.2, 1, 0.3, 0.3, 0.3, 0.3,
0.2, 0.2, 0.2, 0.2, 0.2, 0.3, 1, 0.3, 0.3, 0.3,
0.2, 0.2, 0.2, 0.2, 0.2, 0.3, 0.3, 1, 0.3, 0.3,
0.2, 0.2, 0.2, 0.2, 0.2, 0.3, 0.3, 0.3, 1, 0.3,
0.2, 0.2, 0.2, 0.2, 0.2, 0.3, 0.3, 0.3, 0.3, 1
), ncol=10)

source("http://aoki2.si.gunma-u.ac.jp/R/src/gendat.R", encoding="euc-jp")
z <- gendat(120, r)   # 120 でサンプルサイズを指定

i1 <- bin.var(z[,1], bins=5, method='intervals', labels=c('1','2', '3','4','5')) #  数値変数を区間に分ける
i1 <- as.numeric(as.character(i1)) #数値変数に戻す
i2 <- bin.var(z[,2], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i2 <- as.numeric(as.character(i2))
```

```
i3 <- bin.var(z[,3], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i3 <- as.numeric(as.character(i3))
i4 <- bin.var(z[,4], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i4 <- as.numeric(as.character(i4))
i5 <- bin.var(z[,5], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i5 <- as.numeric(as.character(i5))
i6 <- bin.var(z[,6], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i6 <- as.numeric(as.character(i6))
i7 <- bin.var(z[,7], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i7 <- as.numeric(as.character(i7))
i8 <- bin.var(z[,8], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i8 <- as.numeric(as.character(i8))
i9 <- bin.var(z[,9], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i9 <- as.numeric(as.character(i9))
i10 <- bin.var(z[,10], bins=5, method='intervals', labels=c('1','2', '3','4','5'))
i10 <- as.numeric(as.character(i10))

x <- data.frame(i1, i2, i3, i4, i5)
y <- data.frame(i6, i7, i8, i9, i10)

alphax[i] <- cronbach.alpha(x)[1]
alphay[i] <- cronbach.alpha(y)[1]

f1 <- (i1 + i2 + i3 + i4 + i5)/5
f2 <- (i6 + i7 + i8 + i9 + i10)/5

correl[i] <- cor(f1, f2)
suppressWarnings(rx <- as.numeric(cronbach.alpha(x)))
suppressWarnings(ry <- as.numeric(cronbach.alpha(y)))
adjust[i] <- cor(f1, f2)/sqrt(rx[1]*ry[1])    # Correction for attenuation

}

# 平均を求める
mean(as.numeric(alphax))
mean(as.numeric(alphay))
mean(correl)
mean(adjust)




# 検定力分析
library(pwr)
pwr.r.test(r=0.5, sig.level = 0.05, power = 0.80)
```

```
# SEM
library(sem)　#sem パッケージ使用

model <- specify.model()
i1 < f1, NA, 1
i2 < f1, path2, NA
i3 < f1, path3, NA
i4 < f1, path4, NA
i5 < f1, path5, NA
i6 < f2, NA, 1
i7 < f2, path7, NA
i8 < f2, path8, NA
i9 < f2, path9, NA
i10 < f2, path10, NA
f1 <> f2, path11, NA
i1 <> i1, e1, NA
i2 <> i2, e2, NA
i3 <> i3, e3, NA
i4 <> i4, e4, NA
i5 <> i5, e5, NA
i6 <> i6, e6, NA
i7 <> i7, e7, NA
i8 <> i8, e8, NA
i9 <> i9, e9, NA
i10 <> i10, e10, NA
f1 <> f1, e11, NA
f2 <> f2, e12, NA


ans1 <- sem(model, r, N=N)
summary(ans1)

# 標準化係数
std.coef(ans1)

# パス図作成
path.diagram(ans1, "ans1", ignore.double=FALSE, edge.labels="values", standardize=TRUE, same.rank="f1, f2")
```