

# 無効回答の発生

田中 重人

(大阪大学人間科学部)

Invalid Answers

TANAKA Sigeto

全国家族調査 (NFR98) の質問項目のうち、非該当ケースの存在しない 50 項目について、無効回答 (無回答および「その他」「わからない」) の発生状況を確認する。主な結果はつぎのとおり：(1) 全 6985 ケース中 3214 ケース (46%) で無効回答が生じている；(2) 回収率の高い調査地域で無効回答の多い場合がある；(3) 初等・中等学歴の高齢者、非有職者、無配偶者に無効回答が多い；(4) 「その他」「わからない」を回答選択肢にふくむ項目、数字を直接書き込む項目は無効回答が多い；(5) 調査員が回収時にチェックする項目は無効回答がすくない；(6) 同種の質問項目がワン・セットになっている設問の場合、各項目の無効回答発生はたがいに強く関連している；(7) 基本的属性が無効回答にあたえる影響は項目によってちがうが、学歴の効果は 50 項目すべてに一貫してみられた。これらの結果をもとに、分析上注意すべき点と今後の調査で改善すべき点にも言及する。

キーワード：欠損データ, 無回答, 非標本誤差

## 1. 問題の所在

本稿の目的は、NFR98 データのなかの無効回答について基礎的な分析をおこなうことである。無回答などの欠損データをふくむ変数は分析結果にゆがみをもたらすことがあるため、慎重な取り扱いが要求される。特に NFR データはつぎのような特徴をそなえているので、無効な回答の分布やその原因について基礎的な知識を共有しておくことのメリットが大きい。

### 1. 大勢の研究者が共有するデータである。

NFR データは約 70 人という大規模な研究チームが共同で利用するものであり、また報告書刊行後に一般公開の予定となっている (日本家族社会学会, 2000: まえがき, 1)。これが小規模なプロジェクトなら、データのゆがみについての注意事項を

メンバーの間で口伝で共有していくこともできるだろう。だがこれだけ大人数になるとそれはむずかしい。やはりきちんとした文書のかたちで知識を書きのこしておく必要がある。そうした文書があれば、実査やデータセット構築の過程を知らない研究者がデータ公開後に新しく参入してくる場合にも、必要な事項を確実に伝えることができる。

## 2. 継続・反復調査である。

今回のNFR98調査は継続的調査の第1回目という位置づけであり、今後同様の調査をくりかえしおこなう予定になっている(日本家族社会学会, 2000: 3)。調査法上の工夫で無効回答によるデータのゆがみを小さくすることができる見通しがたてば、今後の調査の精度の向上が期待できる。

## 2. 方法

### (1) 項目の限定

分析に使うのはNFR98 ver.2 (2000年4月14日配布) データである(日本家族社会学会, 2000: 109)。つぎの手順で、非該当の存在しない項目だけを選んだ:

1. NFR98 ver.2 データ付属のSPSSシンタックス *nfrver2spssprogram.doc* ファイルで missing values が (9) または (99) と指定されている変数をえらんだ。
2. これらの変数のうち、Q07G, Q21{1,2,3,6,7}, Q30 は、実際には非該当が存在するので、分析からのぞいた。

結局、分析に使うのは50項目である(表1)。

### (2) 「無効回答」の定義

表1の50項目をそれぞれ2値変数(有効回答=1, 無効回答=0)にコードしなおして分析する。ただしここで「無効回答」としているのはつぎのふたつの場合をふくむ。

無回答 調査票の選択肢に丸がついていないか、記入欄が空のままになっている場合。データ・ファイル中では「9」または「99」の数値が割り振られている。

その他・わからない 学歴(問4)、父親の学歴・職業(問5, 6)、住居形態(問14)、世帯年収(問15)の各項目には「その他」または「わからない」という選択肢がある。これらは分析のうえでは欠損値としてあつかわれることが多いと判断して、無効回答にふくめることにした<sup>1)</sup>。

そして「無回答」「その他」「わからない」のいずれでもない場合を「有効回答」とする。

表-1 変数一覧

変数	質問項目	p.
Q03	小学校卒業まで一番長く暮らしていた地域	1
Q0401	学歴	2
Q05	父親の学歴	
Q06	父親の職業	
Q07	親元を離れて暮らした有無	
Q08	就業状況	3
Q10	本人年収	5
Q11	同居人数	
Q13	世帯主	
Q14	住居形態	
Q15	世帯年収	
Q16	未既婚	6
Q2001	意見(ア)男性は外で働き、女性は家庭を守るべきである	10
Q2002	意見(イ)愛のない夫婦は離婚すべきだ	
Q2003	意見(ウ)未婚者でもお互いに強い愛情があれば性的な関係をもってもかまわない	
Q2004	意見(エ)子供のためなら、親は自分のことを犠牲にすべきだ	
Q2005	意見(オ)親の面倒を見るのは長男の義務である	
Q2006	意見(カ)親が年をとって自分たちで暮らせなくなったら子どもは親と同居すべき	
Q2104	この1ヵ月ほどの間に(エ)「自分が家族に理解されていない」と感じた	11
Q2105	この1ヵ月ほどの間に(オ)家族内での自分の負担が大きすぎると感じた	
Q22	この一年間の健康状態	
Q2301	この1週間の心身状態(ア)ふだんは何でもないことをわずらわしいと感じたこと	
Q2302	この1週間の心身状態(イ)家族や友達から励ましてもらっても気分が晴れないこと	
Q2303	この1週間の心身状態(ウ)憂うつだと感じたこと	
Q2304	この1週間の心身状態(エ)他の人と同じ程度の能力があると思ったこと	
Q2305	この1週間の心身状態(オ)物事に集中できなかったこと	
Q2306	この1週間の心身状態(カ)食欲が落ちたこと	
Q2307	この1週間の心身状態(キ)何をしても面倒と感じたこと	
Q2308	この1週間の心身状態(ク)これから先のことについて積極的に考えたこと	
Q2309	この1週間の心身状態(ケ)何か恐ろしい気持ちがあったこと	
Q2310	この1週間の心身状態(コ)なかなか眠れなかったこと	
Q2311	この1週間の心身状態(サ)生活について不満なく過ごせたこと	
Q2312	この1週間の心身状態(シ)ふだんより口数が少なくなったこと	
Q2313	この1週間の心身状態(ス)一人ぼっちで寂しいと感じたこと	
Q2314	この1週間の心身状態(セ)「毎日が楽しい」と感じたこと	
Q2315	この1週間の心身状態(ソ)悲しいと感じたこと	
Q2316	この1週間の心身状態(タ)仕事が手につかなかったこと	
Q24	子どもの人数	12
Q25ABY	父親(ア)出生年	14
Q25AK	父親(イ)健在・死亡の別	
Q25BBY	母親(ア)出生年	
Q25BK	母親(イ)健在・死亡の別	
Q26	きょうだいの人数	15
Q3101A	親族人数(ア)祖父・祖母	19
Q3101B	親族人数(イ)おじ・おば	
Q3101C	親族人数(ウ)甥・姪	
Q3101D	親族人数(エ)子どもの配偶者	
Q3101E	親族人数(オ)孫	
Q3101F	親族人数(カ)兄弟姉妹の配偶者	
Q3101G	親族人数(キ)配偶者の兄弟姉妹	

太線はページ区切り、細線は変数のまとまりをあらわす。右端は調査票の該当ページ番号。

表-2 50項目中の無効回答数の度数分布

有効-無効項目数	50項目中の無効%	度数 (%)	累積度数 (%)
12-38	76%	1 (0.0)	1 (0.0)
15-35	70%	1 (0.0)	2 (0.0)
16-34	68%	1 (0.0)	3 (0.0)
18-32	64%	4 (0.1)	7 (0.1)
19-31	62%	1 (0.0)	8 (0.1)
20-30	60%	2 (0.0)	10 (0.1)
21-29	58%	3 (0.0)	13 (0.2)
22-28	56%	4 (0.1)	17 (0.2)
23-27	54%	7 (0.1)	24 (0.3)
24-26	52%	6 (0.1)	30 (0.4)
25-25	50%	7 (0.1)	37 (0.5)
26-24	48%	6 (0.1)	43 (0.6)
27-23	46%	3 (0.0)	46 (0.7)
28-22	44%	4 (0.1)	50 (0.7)
29-21	42%	9 (0.1)	59 (0.8)
30-20	40%	4 (0.1)	63 (0.9)
31-19	38%	4 (0.1)	67 (1.0)
32-18	36%	10 (0.1)	77 (1.1)
33-17	34%	8 (0.1)	85 (1.2)
34-16	32%	10 (0.1)	95 (1.4)
35-15	30%	7 (0.1)	102 (1.5)
36-14	28%	3 (0.0)	105 (1.5)
37-13	26%	10 (0.1)	115 (1.6)
38-12	24%	4 (0.1)	119 (1.7)
39-11	22%	12 (0.2)	131 (1.9)
40-10	20%	26 (0.4)	157 (2.2)
41- 9	18%	21 (0.3)	178 (2.5)
42- 8	16%	36 (0.5)	214 (3.1)
43- 7	14%	49 (0.7)	263 (3.8)
44- 6	12%	86 (1.2)	349 (5.0)
45- 5	10%	122 (1.7)	471 (6.7)
46- 4	8%	195 (2.8)	666 (9.5)
47- 3	6%	433 (6.2)	1099 (15.7)
48- 2	4%	792 (11.3)	1891 (27.1)
49- 1	2%	1323 (18.9)	3214 (46.0)
50- 0	0%	3771 (54.0)	6985 (100.0)

### 3. 無効回答の発生状況

#### (1) 無効回答の度数分布

分析に使う50項目のなかにどれだけの無効回答がふくまれているか、度数分布表で検討しよう(表2)。全6985ケースのうち、50項目すべてが有効なのは3771ケースである。つまりNFR98データの回収ケースのうち54%は、これら50項目に関しては欠損のない「完全ケース」(保田, 2000: 166)としてあつかえるのだ。のこりの46%(3214ケース)はいずれかの項目に無効回答がある。これらのうち1項目だけが無効であるケースが1323, ついで2項目が無効であるケースが792と、ほとんどのケースは少数の項目での欠損ですんでいる。その一方、5つ以上の項目が欠けているケースも471(6.7%)ある。いちばん

無効回答の多かったケースでは38項目が無効となっており、有効な項目は12個だけしかない。

## (2) 調査地域と無効回答

調査地域（都道府県）別にみた無効回答の発生状況を表3に示す。まず目立って無効回答が多いのは山形県で、完全ケース（無効回答ゼロ）は3割しかおらず、対して5項目以上が無効であるケースが3割をこえている。山形ほどではないが、鳥取・徳島の各県も無効回答が多い。一方、無効回答がすくないのは埼玉・東京・福井・島根・広島などで、いずれも完全ケースが6割をこえている。

図1は、都道府県ごとの回収率を横軸に、無効回答が5項目以上にのぼるケースの比率を縦軸にとってプロットしたものである。回収率が60%を下回っている都府県（千葉・神奈川・京都・大阪・東京・埼玉）では、無効回答5項目以上のケースは10%以下と比較的すくない。これに対して、回収率が比較的高い県のなかには無効回答の多いところもあり、山形・鳥取・徳島・佐賀・長崎の各県では15%以上のケースで5項目以上が無効になっている。47都道府県全体について Kendall の順位相関係数  $\tau$  を計算してみたところ 0.115であった。回収率の高い地域ほど無効回答が多くなるということだ。ただし回収率が高くても無効回答がすくない県（広島・宮崎・福井・島根など）もあるので、回収率と無効回答の間に単純な相関関係があるとはいえない。むしろ、回収率の高い地域のなかに、無効回答のすくないところから多いところまで幅ひろい分化がみられると表現するのが適切だと思う。

## (3) 基本的属性と無効回答

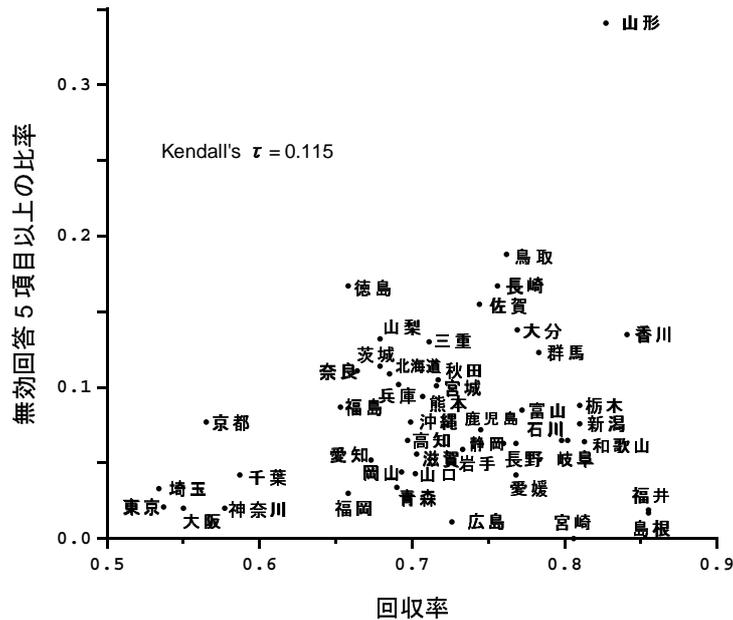
つぎに無効回答数と基本的属性との関連をみておこう。「基本的属性」としてとりあげるのは性別 (Q01)、年齢 (Q02S)、学歴 (Q0401)、就業状況 (Q08)、同居人数 (Q11)、未既婚 (Q16) の6項目である。これらのうち「性別」「年齢」「同居人数」「未既婚」の4項目には無効回答はなく、全ケースが有効である。これに対して「就業状況」「学歴」の2項目では無効回答が生じており、特に学歴では100人以上が無回答となっている（あとででてくる表11を参照）。これら2項目の無効回答は全項目合計の無効回答数に計上されているものであることに留意していただきたい。

表4にこれら基本属性と無効回答数との関連をまとめておく。性別以外の各変数は、無効回答数と関連していることがわかる。個々の属性と無効回答数との関連を読みとると、つぎのことがいえる。

- 年齢が高くなると無効回答が増える。68才以上層では5つ以上の項目に無効回答が生じているケースが13%にも達している。
- 学歴が高くなると無効回答が減る。大学卒以上の学歴層では50項目全部が有効である完全ケースが3/4を占めている。

表-3 調査地域と無効回答発生項目数

都道府県	15以上	5-15	3-4	2	1	0	合計 (N)
北海道	2.1	8.8	8.2	13.4	18.5	48.9	100.0 (329)
青森	1.1	2.3	3.4	13.8	19.5	59.8	100.0 ( 87)
岩手	1.2	4.7	7.1	12.9	23.5	50.6	100.0 ( 85)
宮城	2.2	7.9	11.5	12.2	20.9	45.3	100.0 (139)
秋田	5.3	5.3	13.2	5.3	21.1	50.0	100.0 ( 76)
山形	13.2	20.9	11.0	13.2	9.9	31.9	100.0 ( 91)
福島	0.0	8.7	6.1	13.0	25.2	47.0	100.0 (115)
茨城	2.4	9.0	10.2	10.2	14.4	53.9	100.0 (167)
栃木	1.5	7.4	11.8	15.4	17.6	46.3	100.0 (136)
群馬	1.5	10.8	13.1	10.8	16.9	46.9	100.0 (130)
埼玉	0.3	2.9	6.5	11.4	18.0	60.8	100.0 (306)
千葉	0.0	4.2	5.5	8.3	23.2	58.8	100.0 (289)
東京	0.2	1.9	5.8	9.5	19.6	63.1	100.0 (536)
神奈川	0.0	2.0	8.8	11.5	18.9	58.7	100.0 (407)
新潟	2.3	5.3	10.5	9.9	21.6	50.3	100.0 (171)
富山	5.6	2.8	8.5	8.5	25.4	49.3	100.0 ( 71)
石川	0.0	6.5	18.2	9.1	18.2	48.1	100.0 ( 77)
福井	1.9	0.0	9.4	7.5	20.8	60.4	100.0 ( 53)
山梨	5.7	7.5	15.1	17.0	18.9	35.8	100.0 ( 53)
長野	0.7	5.6	8.5	10.6	21.8	52.8	100.0 (142)
岐阜	0.0	6.5	10.9	10.9	10.9	60.9	100.0 (138)
静岡	1.7	4.6	11.4	13.5	20.3	48.5	100.0 (237)
愛知	1.0	4.2	7.0	11.7	17.5	58.5	100.0 (383)
三重	0.9	12.0	13.0	10.2	18.5	45.4	100.0 (108)
滋賀	1.4	4.2	9.9	19.7	19.7	45.1	100.0 ( 71)
京都	1.7	6.0	8.5	14.5	20.5	48.7	100.0 (117)
大阪	0.2	1.7	8.7	10.9	19.4	59.0	100.0 (402)
兵庫	4.4	5.7	9.8	8.6	22.5	48.9	100.0 (315)
奈良	0.0	11.1	13.6	8.6	19.8	46.9	100.0 ( 81)
和歌山	1.3	5.1	5.1	9.0	16.7	62.8	100.0 ( 78)
鳥取	3.1	15.6	12.5	6.3	21.9	40.6	100.0 ( 32)
島根	1.7	0.0	15.3	10.2	10.2	62.7	100.0 ( 59)
岡山	0.9	3.5	11.5	8.8	20.4	54.9	100.0 (113)
広島	0.0	1.1	5.1	13.1	15.4	65.1	100.0 (175)
山口	1.1	3.3	12.0	10.9	19.6	53.3	100.0 ( 92)
徳島	0.0	16.7	16.7	6.3	16.7	43.8	100.0 ( 48)
香川	5.4	8.1	5.4	10.8	17.6	52.7	100.0 ( 74)
愛媛	0.0	4.2	13.5	10.4	15.6	56.3	100.0 ( 96)
高知	2.2	4.3	17.4	13.0	19.6	43.5	100.0 ( 46)
福岡	0.4	2.6	6.0	12.1	20.4	58.5	100.0 (265)
佐賀	1.7	13.8	3.4	13.8	20.7	46.6	100.0 ( 58)
長崎	6.3	10.4	11.5	12.5	11.5	47.9	100.0 ( 96)
熊本	0.9	8.5	7.5	9.4	18.9	54.7	100.0 (106)
大分	1.3	12.5	15.0	20.0	11.3	40.0	100.0 ( 80)
宮崎	0.0	0.0	3.8	10.1	26.6	59.5	100.0 ( 79)
鹿児島	1.8	5.4	9.0	15.3	12.6	55.9	100.0 (111)
沖縄	3.1	4.6	16.9	15.4	21.5	38.5	100.0 ( 65)
合計 (N)	1.5 (102)	5.3 (369)	9.0 (628)	11.3 (792)	18.9 (1323)	54.0 (3771)	100.0 (6985)



回収率は日本家族社会学会 [2000:12] による。

図-1 都道府県ごとの回収率と無効回答

- 仕事に就いていると無効回答が減る。
- 単身生活者は無効回答が多い。逆に4-5人の世帯では無効回答が少ない。
- 配偶者と離別/死別している場合に無効回答が多い。

NFR98調査の最初の報告書(日本家族社会学会, 2000: 12-16)には「国勢調査」などの全国規模データの属性の分布とNFR98回収票とのずれの検討結果がのっている。それによると、無配偶の者、単独世帯の者、義務教育修了者の回収がすくなく、そのぶん有配偶者、親族世帯、中・高等学歴者の構成比が高くなっているとのことである。これらの属性による回収状況の差は、無効回答の発生状況と部分的に重なる。無配偶者・単独世帯・義務教育修了者は無効回答が多いため分析上欠損ケースとしてはぶかれる可能性が高いのだが、それにくわえて回収率がそもそも低い。これらの要因によって、NFR98データによる分析結果は、中等以上の学校を出て、配偶者がいて、親族と一緒に住んでいる層にかたよったものになりやすいといえる。

#### (4) 属性変数の組み合わせと無効回答

ただしこれらの属性変数はたがいに関連しているから、表4の解釈には慎重さが必要である。たとえば高学歴化の結果として、年齢と学歴の間には負の相関関係がある。高学歴層の無効回答がすくないのは、単に若い人が多いためかもしれない。また、若年層と高齢

表-4 基本的属性と無効回答発生項目数

	15以上	5-15	3-4	2	1	0	合計 (人数)
Q01(性別)							$\chi^2=9.647, df=5, p>0.05.$
男性	1.4	4.9	8.6	11.5	17.9	55.7	100.0 (3323)
女性	1.5	5.6	9.3	11.2	19.9	52.5	100.0 (3662)
Q02S(年齢)							$\chi^2=673.644, df=20, p<0.05.$
28-37才	0.6	2.1	3.9	7.7	20.6	65.0	100.0 (1410)
38-47才	1.0	3.1	4.5	7.0	18.0	66.4	100.0 (1453)
48-57才	0.7	5.0	8.2	11.0	20.4	54.8	100.0 (1668)
58-67才	2.1	7.2	11.8	13.6	19.6	45.6	100.0 (1353)
68-77才	3.5	10.3	19.1	19.4	15.2	32.5	100.0 (1101)
Q0401(学歴)							$\chi^2=1211.981, df=30, p<0.05.$
中学校 (旧制小学校)	2.9	9.3	17.5	18.2	19.3	32.8	100.0 (1690)
高校 (旧制中学・高女)	0.9	4.3	7.3	10.2	19.9	57.4	100.0 (2910)
専門学校	0.5	1.6	5.4	9.1	22.8	60.6	100.0 (373)
短大・高専 (旧制高校・専門)	0.4	2.3	4.6	7.7	18.7	66.4	100.0 (791)
大学以上	0.5	1.9	3.0	5.3	14.3	75.0	100.0 (1061)
その他	0.0	7.3	9.8	31.7	51.2	—	100.0 ( 41)
無回答	13.4	33.6	22.7	19.3	10.9	—	100.0 (119)
Q08(就業状況)							$\chi^2=246.091, df=20, p<0.05.$
就いている	1.0	3.9	7.2	9.7	19.0	59.2	100.0 (4595)
休職中	1.1	10.2	11.4	13.6	21.6	42.0	100.0 ( 88)
無職 (経験あり)	2.3	7.4	12.3	14.6	18.7	44.7	100.0 (2098)
仕事経験なし	2.5	12.0	14.5	13.0	20.5	37.5	100.0 (200)
無回答	0.0	50.0	25.0	25.0	0.0	—	100.0 ( 4)
Q11(同居人数)							$\chi^2=169.839, df=25, p<0.05.$
1人	2.8	11.1	16.5	16.0	14.2	39.3	100.0 (351)
2人	1.2	5.5	11.7	13.0	18.0	50.6	100.0 (1698)
3人	1.5	4.9	7.2	11.4	20.0	54.9	100.0 (1599)
4人	1.1	4.0	5.7	9.4	19.7	60.2	100.0 (1502)
5人	0.9	4.9	8.2	8.9	18.5	58.6	100.0 (901)
6人以上	2.5	5.8	10.4	12.0	19.8	49.6	100.0 (934)
Q16(未既婚)							$\chi^2=198.493, df=10, p<0.05.$
有配偶	1.2	4.8	8.4	10.6	19.0	56.0	100.0 (5677)
離死別	3.8	9.5	17.3	17.3	18.5	33.6	100.0 (676)
未婚	1.3	4.7	5.7	11.9	19.0	57.4	100.0 (632)
合計	1.5	5.3	9.0	11.3	18.9	54.0	100.0
(人数)	(102)	(369)	(628)	(792)	(1323)	(3771)	(6985)

層とでは学歴の影響のしかたがちがう、というような交互作用効果があるかもしれない。そこでいくつかの変数の組み合わせで無効回答の出現頻度をチェックしてみた。

表5は、5項目以上が無効になっているケースの比率が学歴と年齢の組み合わせでどうかわるかを示している。まず、いずれの年齢層でも、学歴があがるにしたがって無効回答が減るという傾向があることを確認しておこう。また「中学」(義務教育卒)、「高校」(中等教育卒)では年齢の効果が大きくあらわれている。特に「中学」で58才以上の年齢層では5項目以上無効ケースが10%をこえている。学歴が無回答の場合も、年齢があがる

表-5 学歴 (Q0401)・年齢 (Q02S) と5項目以上の無効回答発生率

年齢	中学	高校	専門	短大	大学	その他	無回答
28-37	11.9 (42)	2.5 (594)	1.6 (126)	2.6 (311)	2.5 (321)	12.5 (8)	0.0 (8)
38-47	7.8 (115)	4.5 (665)	3.0 (100)	3.0 (231)	2.2 (323)	0.0 (12)	42.9 (7)
48-57	9.0 (390)	4.7 (788)	2.1 (94)	3.1 (127)	2.2 (229)	0.0 (13)	40.7 (27)
58-67	11.8 (550)	7.9 (529)	2.2 (46)	1.7 (59)	2.2 (138)	50.0 (4)	48.1 (27)
68-77	15.5 (593)	8.4 (334)	0.0 (7)	1.6 (63)	4.0 (50)	0.0 (4)	58.0 (50)
年齢計	12.2 (1690)	5.2 (2910)	2.1 (373)	2.7 (791)	2.4 (1061)	7.3 (41)	47.1 (119)

% (人数) を示す。

表-6 就業状況 (Q08)・年齢 (Q02S) と5項目以上の無効回答発生率

年齢	有職	休職	無職	未経験	無回答
28-37	2.7 (1083)	0.0 (18)	3.0 (303)	16.7 (6)	— (0)
38-47	3.8 (1220)	11.1 (18)	3.0 (201)	35.7 (14)	— (0)
48-57	5.4 (1353)	5.3 (19)	7.0 (272)	4.2 (24)	— (0)
58-67	7.8 (703)	18.8 (16)	11.2 (565)	7.5 (67)	50.0 (2)
68-77	9.3 (236)	23.5 (17)	14.3 (757)	19.1 (89)	50.0 (2)
年齢計	4.9 (4595)	11.4 (88)	9.8 (2098)	14.5 (200)	50.0 (4)

% (人数) を示す。

表-7 未既婚 (Q16)・同居人数 (Q11)・年齢 (Q02S) と5項目以上の無効回答発生率

年齢	有配偶		離死別			未婚	
	単独	複数	単独	2人	3人以上	単独	複数
28-37	33.3 (3)	2.4 (968)	0.0 (12)	10.0 (10)	3.4 (29)	3.2 (63)	3.4 (325)
38-47	0.0 (1)	3.9 (1252)	10.0 (10)	5.0 (20)	2.7 (37)	6.9 (29)	4.8 (104)
48-57	16.7 (6)	5.2 (1473)	13.8 (29)	2.2 (45)	7.5 (53)	20.0 (25)	8.1 (37)
58-67	0.0 (4)	8.7 (1159)	8.2 (49)	6.1 (49)	21.5 (65)	17.6 (17)	20.0 (10)
68-77	0.0 (1)	11.4 (810)	26.5 (83)	11.3 (53)	20.5 (132)	21.1 (19)	33.3 (3)
年齢計	13.3 (15)	6.0 (5662)	16.9 (183)	6.8 (177)	14.9 (316)	10.5 (153)	4.6 (479)

% (人数) を示す。

にしたがって無効回答がふえる傾向がある。しかし高等教育卒業業者である「短大」「大学」になると、このような年齢の効果はみられず、どの年齢層でも5項目以上無効ケースは4%以下である。つまり学歴と年齢は無効回答の発生に対して交互作用効果をおよぼしており、初等・中等学歴の高年齢者に特に無効回答が多い。

表6は就業状況と年齢を組み合わせしてみた無効回答の発生状況である。職業についての経験がない場合、若年層と高年齢層で無効回答が多く、中間の年齢層(48-67才)で比較的無効回答が少ないというU字型の関連になっている。それ以外の「休職」「無職」「有職」の各層では、年齢があがるにしたがって無効回答が直線的に増えていく。

表7は未既婚と世帯人数とを組み合わせ7つの類型を設定し、それらの類型と年齢

とで5項目以上無効回答ケースの比率がどうちがうかを示している。有配偶の場合にくらべて未婚・離死別の場合に無効回答が多く、その傾向は年齢があがるにしたがって強くなる。ただしどういうわけか、離死別でふたり暮らしの場合は、有配偶者とあまりパターンがかわらない。

表-8 無効回答数5以上の発生を予測するロジスティック回帰

独立変数	推定値 (標準誤差)		推定値 (標準誤差)	
(定数項)	-4.152*	(0.260)	-2.265*	(0.405)
性別 [男=1, 女=2]	0.056	(0.100)	-0.198	(0.114)
年齢 [基準=28-37才]				
38-47	0.493*	(0.216)	0.366	(0.218)
48-57	0.903*	(0.205)	0.463*	(0.214)
58-67	1.486*	(0.204)	0.787*	(0.221)
68-77	1.839*	(0.205)	0.829*	(0.235)
同居人数 (本人ふくむ) [基準=2人]				
1人	0.598*	(0.214)	0.515*	(0.229)
3人	0.318*	(0.146)	0.292	(0.155)
4人	0.312	(0.163)	0.251	(0.174)
5人	0.314	(0.181)	0.234	(0.193)
6人以上	0.510*	(0.158)	0.395*	(0.168)
未既婚 [基準=有配偶]				
離死別	0.410*	(0.152)	0.446*	(0.160)
未婚	0.498*	(0.211)	0.462*	(0.219)
学歴 [基準=中学校]				
高校・専門学校			-0.660*	(0.121)
短大・高専			-1.166*	(0.247)
大学以上			-1.257*	(0.221)
就業状況 [基準=就業経験なし]				
休職中・現在無職			-0.335	(0.233)
有職			-0.665*	(0.249)
標本数	6985		6863	
うち5項目以上無効	471		414	
$-2 \log L$	3262.335		2913.884	
$\chi^2$	187.422*		213.676*	
DF	12		17	

\* :  $p < 0.05$ .

これらの属性変数を同時に投入したロジスティック回帰分析 (SAS, 1990) の結果を表8に示す。交互作用項をふくまない単純な加法モデルである。ただし学歴と就業状況は無効回答をふくんでいるので、これら以外の変数だけによるモデルと、それにこれらを加えたモデルの2つを推定することにした。

性別・年齢・同居人数・未既婚を投入したモデルの推定結果 (表8左側) をみると、性別以外の各変数が有意な効果を持っていることがわかる。いずれの変数についても、単純なクロス表 (表4) で検討したのとおなじ方向での効果がみられる。学歴と就業状況をくわえたモデルの推定結果 (表8右側) でもこれらの傾向は同様である。ただし年齢の効果はかなり小さくなり、各年齢カテゴリーの係数はほぼ半分の大きさに落ちている。すでに

表-9 1項目以上の無効回答の発生を予測するロジスティック回帰

独立変数	推定値 (標準誤差)		推定値 (標準誤差)	
(定数項)	-1.034*	(0.115)	0.374	(0.224)
性別 [男=1, 女=2]	0.093	(0.050)	-0.057	(0.057)
年齢 [基準=28-37才]				
38-47	-0.018	(0.081)	-0.065	(0.083)
48-57	0.510*	(0.079)	0.247*	(0.083)
58-67	0.881*	(0.086)	0.404*	(0.094)
68-77	1.356*	(0.094)	0.681*	(0.107)
同居人数 (本人ふくむ) [基準=2人]				
1人	0.261	(0.138)	0.261	(0.142)
3人	0.140	(0.074)	0.139	(0.077)
4人	0.132	(0.079)	0.141	(0.082)
5人	0.139	(0.091)	0.117	(0.094)
6人以上	0.367*	(0.087)	0.281*	(0.090)
未婚 [基準=有配偶]				
離死別	0.603*	(0.097)	0.549*	(0.099)
未婚	0.316*	(0.100)	0.298*	(0.102)
学歴 [基準=中学校]				
高校・専門学校			-0.754*	(0.068)
短大・高専			-1.045*	(0.100)
大学以上			-1.303*	(0.093)
就業状況 [基準=就業経験なし]				
休職中・現在無職			-0.104	(0.162)
有職			-0.309	(0.165)
標本数	6985		6863	
うち無効回答あり	3214		3092	
$-2 \log L$	9146.986		8740.537	
$\chi^2$	491.817*		706.313*	
DF	12		17	

係数の推定値 (標準誤差) を示す。\* :  $p < 0.05$ .

表5で検討したように、年齢の効果は学歴の効果にかなりの部分吸収されるのだ。

ここまでの分析では5項目以上が無効になっているかどうか注目してきた。同様の分析を、すくなくとも1項目が無効となっているかどうかについてもおこなっている。結果はいちいち示さないが、5項目以上無効の場合と似た結果となった。表8とおなじ変数を投入したロジスティック回帰分析の結果を表9に示す。

ただし学歴と年齢との交互作用に関しては結果がちがっていたので、それだけはクロス表を示しておく。表10に示したとおり、1項目以上の無効回答が生じているかどうか、という観点からみると、学歴と年齢との交互作用効果はみられない——学歴の高低にかかわらず、年齢が高くなるほど無効回答の発生率が高くなっている。

## (5) まとめ

以上の分析から NFR98 データの無効回答と基本的属性との関連をまとめておこう。

表-10 学歴 (Q0401)・年齢 (Q02S) と1項目以上の無効回答発生率

年齢	中学	高校	専門	短大	大学
28-37	50.0 (42)	39.2 (594)	38.1 (126)	31.5 (311)	24.0 (321)
38-47	51.3 (115)	36.2 (665)	31.0 (100)	29.9 (231)	21.4 (323)
48-57	60.3 (390)	42.3 (788)	44.7 (94)	37.0 (127)	24.9 (229)
58-67	67.6 (550)	46.3 (529)	52.2 (46)	40.7 (59)	29.0 (138)
68-77	75.7 (593)	56.3 (334)	28.6 (7)	44.4 (63)	44.0 (50)
年齢計	67.2 (1690)	42.6 (2910)	39.4 (373)	33.6 (791)	25.0 (1061)

% (人数) を示す。

- 年齢が高くなると無効回答が増える。ただしこの効果の半分程度は学歴によって説明できる。
- 学歴が高くなると無効回答は減る。
- 学歴と年齢の間には交互作用がみられる。学歴の低い層ほど、年齢による無効回答増加傾向が顕著にあらわれる。
- 就業経験と年齢とは無効回答の発生に交互作用的に影響をおよぼしている。
- 配偶者がいない高齢者では無効回答が多い。
- 性別は無効回答の発生に影響しない。

#### 4. 項目別無効回答の分析

##### (1) 項目別にみた無効回答の発生状況

つぎに、各項目で無効回答の発生にどのようなちがいがあがあるかをみていこう。表11は質問項目ごとに無効回答ケースを数え、すくない順にならべかえたものである。無効回答が5%を上回っているのは、父母の出生年、父の学歴・職業、世帯年収の各項目である。特に父母の出生年は無効回答が12-15%と際立って多い。これらの項目を使って分析をする場合は要注意である。

2.(2)でことわっておいたように、本稿では無回答のほかに「その他」「わからない」といった回答も「無効回答」としてあつかう。たとえば父親の職業 (Q06) は無回答ケースは51しかないのだが、「その他」が48ケース、「わからない」が255ケースあるために、結局354ケースが無効になっている。本人学歴 (Q0401)、父親学歴 (Q05)、住居形態 (Q14)、世帯年収 (Q15) も同様の事情で、無回答以外の理由による無効ケースが出ている。これらの項目については、表11の「質問内容」の列に [] でくくって無効原因別の内訳人数を示した。

表11から、「その他」「わからない」といった選択肢が無効回答の増加に貢献していることがわかる。無効回答の多い上位5項目のうち3つ (Q06, Q15, Q05) はこれらの選択肢が設定されていた項目であり、しかも無回答ケースよりも「わからない」と答えたケー

表-11 質問項目別にみた無効回答発生状況

変数名	無効回答者数	無効回答%	有効回答%	質問内容
Q11*#	0	0.00	100.00	同居人数
Q16†	0	0.00	100.00	未既婚
Q08†	4	0.06	99.94	就業状況
Q24†#	9	0.13	99.87	子どもの人数
Q13*	13	0.19	99.81	世帯主
Q26†#	20	0.29	99.71	きょうだいの人数
Q14	21	0.30	99.70	住居形態 [無回答 8+その他 13]
Q03	24	0.34	99.66	小学校卒業まで一番長く暮らしていた地域
Q3101D#	31	0.44	99.56	親族人数(エ)子どもの配偶者
Q25BK	32	0.46	99.54	母親(イ)健在・死亡の別
Q25AK	60	0.86	99.14	父親(イ)健在・死亡の別
Q3101E#	86	1.23	98.77	親族人数(オ)孫
Q2001	95	1.36	98.64	意見(ア)男性は外で働き女性は家庭を守る
Q07	98	1.40	98.60	親元を離れて暮らした有無
Q2006	100	1.43	98.57	意見(カ)親が年をとったら子どもは親と同居す
Q2301	106	1.52	98.48	身心状態(ア)ふだんは何でもないことを煩
Q2310	107	1.53	98.47	身心状態(コ)なかなか眠れなかった
Q2306	111	1.59	98.41	身心状態(カ)食欲が落ちた
Q2307	120	1.72	98.28	身心状態(キ)何をしても面倒と感じた
Q2005	127	1.82	98.18	意見(オ)親の面倒を見るのは長男の義務
Q2303	128	1.83	98.17	身心状態(ウ)憂うつだと感じた
Q2315	131	1.88	98.12	身心状態(ソ)悲しいと感じた
Q2313	132	1.89	98.11	身心状態(ス)一人ぼっちで寂しいと感じた
Q2309	133	1.90	98.10	身心状態(ケ)何か恐ろしい気持ちがあった
Q2316	135	1.93	98.07	身心状態(タ)仕事が手につかなかった
Q2305	137	1.96	98.04	身心状態(オ)物事のに集中できなかった
Q2312	141	2.02	97.98	身心状態(シ)ふだんより口数が少なくなっ
Q10	143	2.05	97.95	本人年収
Q2308	143	2.05	97.95	身心状態(ク)これから先の事について積極
Q2004	144	2.06	97.94	意見(エ)子供のためなら親は自分の事を犠牲に
Q2002	158	2.26	97.74	意見(イ)愛のない夫婦は離婚すべき
Q0401*	160	2.29	97.71	学歴 [無回答 119+その他 41]
Q2302	169	2.42	97.58	身心状態(イ)励ましてもらっても気分が晴
Q2311	175	2.51	97.49	身心状態(サ)生活について不満なく過ごせ
Q2314	177	2.53	97.47	身心状態(セ)「毎日が楽しい」と感じた
Q2003	182	2.61	97.39	意見(ウ)未婚者でも性的な関係を持ってもかま
Q22	192	2.75	97.25	この一年間の健康状態
Q2105	204	2.92	97.08	この1月間(オ)家族内の自分の負担大きすぎ
Q3101F#	223	3.19	96.81	親族人数(カ)兄弟姉妹の配偶者
Q2104	228	3.26	96.74	この1月間(エ)「家族に理解されていない」
Q3101A#	244	3.49	96.51	親族人数(ア)祖父・祖母
Q3101G#	267	3.82	96.18	親族人数(キ)配偶者の兄弟姉妹
Q3101C#	290	4.15	95.85	親族人数(ウ)甥・姪
Q3101B#	329	4.71	95.29	親族人数(イ)おじ・おば
Q2304	348	4.98	95.02	身心状態(エ)他の人と同じ程度の能力があ
Q06	354	5.07	94.93	父親の職業 [無回答 51+その他 48+わからない 255]
Q15	679	9.72	90.28	世帯年収 [無回答 130+わからない 549]
Q25BBY#	897	12.84	87.16	母親(ア)出生年
Q05	965	13.82	86.18	父親の学歴 [無回答 96+その他 20+わからない 849]
Q25ABY#	1044	14.95	85.05	父親(ア)出生年

無効回答のすくない順にならべた。N=6985。無回答以外の無効原因がある場合 [] 内に内訳を示す。  
 変数名右肩の印はNFR98 調査要領「D. 調査票の点検事項」(日本家族社会学会, 1999: 56-57) 記載項目  
 (\*: その項目がチェック対象; †: 枝問がチェック対象) および回答形式 (#: 数字記入) をあらわす。

スのほうがはるかに多い。

また数字をマスの中に書いてもらう回答形式も無効回答の原因になっている可能性がある。無効回答の多い上位 10 項目をとりあげてみると、うち 6 つはこの数字書き込み式の項目 (表 11 の #印)、3 つは上記の「その他」「わからない」を選択肢にふくむ項目であり、どちらにも該当しないのは Q2304 だけである。

つぎに無効回答を抑制する要因について考察しておこう。NFR98 調査は留置法による調査であり、回収に際して調査員が調査票のチェックをおこなっている。表 11 からは、このチェックが無効回答を抑制していることが読み取れる。表 11 の変数名の後についているマークは「調査員の手引き」(調査要領)に「回収時に点検すべき項目」としてあがっていた項目を示している。\* マークはその項目自体が点検対象だったもの、† マークはその項目自体は点検対象ではなかったが下位項目が点検対象だったものである。無効回答のすくない項目の 6 位まで (Q11, Q16, Q08, Q24, Q13, Q26) は、これら 2 種類の点検項目に入っていた質問が占めている。

ただし回収時の点検は絶対的な効果があるわけでもない。Q0401 (本人の学歴) は点検項目にあがっていたにもかかわらず、かなり多くの無回答や「その他」が出たため、有効回答は 97.71%にとどまっている。

## (2) 無効回答の布置

これらの項目の無効回答発生状況はたがいにもどういう関係にあるだろうか。分析に使ってきた 50 個の変数のうち、無効回答のすくない 11 個の変数 (表 11 の Q25AK より上) をのぞいて、のこる 39 の変数だけにしぼって分析しよう。

まず 2 値変数同士の関連をあらわす四分点相関係数  $\phi$  を求めてみる。表 12 は 39 変数について  $\phi$  (を 100 倍した値) を行列のかたちで表記したものである。

表 12 は、これら 39 変数中に結びつきの強い変数群がいくつかあることを示している。

まず問 20, 21, 23, 25, 31 のようなワン・セット型の設問は、それぞれの内部で強い相関を持っている。問 20 (家族に関する意見) を構成する 6 変数間の相関はすべて  $\phi=0.5$  以上である。問 21 (ここ 1 ヶ月の悩み) のふたつの変数間には  $\phi=0.78$  の強い関連がある。問 23 (心身状態) を構成する 16 の変数間には、ひとつだけ  $\phi=0.37$  の組み合わせがあるが、ほかはすべて  $\phi=0.4$  を上回っている。問 25 (父母の出生年) のふたつの変数間の相関は  $\phi=0.83$ 。問 31 (親族構成) の 6 変数間はずこし相関が弱く  $0.29 \leq \phi \leq 0.62$  である。

また、ひろい意味で「意識」をたずねている問 20, 21, 23 の諸変数はたがいにもほぼ  $\phi=0.3$  以上の相関を持っており、これらの項目の無回答の発生メカニズムに共通点があることをうかがわせている。

これら以外の変数間にはほとんど関連がみられない。ただし収入に関する変数間 (問 10 と問 15) と親に関する変数間 (問 5, 問 6, 問 25) でそれぞれ  $\phi=0.2$  をこえる相関がいく



つかあるので、それらについて検討しておこう。

表-13 世帯年収への回答と本人年収無回答率

	世帯年収 (Q15)		
	有効回答	わからない	無回答
本人年収 (Q10) 無回答率 (人数)	0.5 (6306)	7.3 (549)	55.4 (130)

% (人数) を示す。

表13では、世帯年収が無回答であるケースの55.4%が本人年収も無回答であることがわかる。世帯年収が「わからない」だった場合も、本人年収の無回答は7%あり、やや高い割合となっている。

表-14 父学歴への回答と父職業への回答

父学歴 (Q05)	父職業 (Q06)				合計 (人数)
	有効回答	その他	わからない	無回答	
有効回答	96.9	0.5	2.1	0.4	100.0 (6020)
その他	95.0	0.0	5.0	0.0	100.0 (20)
わからない	83.0	1.6	14.6	0.7	100.0 (849)
無回答	75.0	1.0	4.2	19.8	100.0 (96)
合計	94.9	0.7	3.7	0.7	100.0 (6985)

父親の学歴と職業というふたつの項目間の関係をみた表14によると、どちらも「わからない」、あるいはどちらも無回答という組み合わせの多いことがわかる。父学歴が有効回答であった場合の父職業の無効回答は3%しかないのに対して、父学歴が「わからない」や無回答であった場合には、父職業の無効回答が2割にも達している。

表-15 父学歴への回答と父出生年無回答率

	父学歴 (Q05)			
	有効回答	その他	わからない	無回答
父出生年 (Q25ABY) 無回答率 (人数)	12.1 (6020)	20.0 (20)	31.9 (849)	43.8 (96)

% (人数) を示す。

表15では、父学歴と父出生年それぞれへの回答の関連をみている。父出生年 (Q25ABY) は分析に使った50項目のなかでいちばん無回答が多く、約15%が無効であった (表11)。このため全体的に父出生年の無回答率は高いのだが、そのなかでも父学歴が「その他」「わからない」あるいは無回答であった場合、さらに無回答率が高くなっている。

### (3) 項目別の無効回答要因の検討

以下では各項目別に無効回答の発生原因を探っていくことにしよう。

表-16 父母の属性に関する無効回答を予測するロジスティック回帰

独立変数	母出生年 (Q25BBY)		父出生年 (Q25ABY)		父学歴 (Q05)	父職業 (Q06)
(定数項)	-2.458*	(0.328)	-2.247*	(0.316)	-0.811*(0.297)	-1.594*(0.436)
性別 [男=1, 女=2]	-0.166	(0.087)	-0.172*	(0.082)	-0.057 (0.079)	-0.251*(0.124)
年齢 [基準=28-37才]						
38-47	-0.347	(0.213)	0.032	(0.199)	-0.398*(0.125)	-0.517*(0.228)
48-57	-0.052	(0.194)	0.233	(0.193)	-0.452*(0.132)	-0.753*(0.234)
58-67	0.211	(0.203)	0.429*	(0.202)	-0.817*(0.154)	-0.547*(0.250)
68-77	0.792*	(0.211)	1.061*	(0.208)	-0.953*(0.169)	-0.726*(0.270)
同居人数 (本人ふくむ) [基準=2人]						
1人	0.287	(0.190)	0.370*	(0.183)	0.440*(0.170)	0.513*(0.242)
3人	0.166	(0.114)	0.104	(0.107)	0.084 (0.107)	-0.012 (0.166)
4人	0.119	(0.135)	0.047	(0.124)	-0.029 (0.116)	0.015 (0.182)
5人	0.078	(0.148)	0.026	(0.138)	0.011 (0.130)	-0.207 (0.216)
6人以上	0.032	(0.129)	0.055	(0.121)	0.026 (0.124)	0.016 (0.186)
未婚 [基準=有配偶]						
離死別	0.381*	(0.125)	0.329*	(0.119)	0.426*(0.119)	0.285 (0.181)
未婚	-0.308	(0.230)	-0.549*	(0.223)	-0.170 (0.149)	-0.009 (0.241)
学歴 [基準=中学校]						
高校・専門学校	-0.749*	(0.091)	-0.807*	(0.085)	-0.828*(0.089)	-0.969*(0.138)
短大・高専	-1.062*	(0.186)	-1.116*	(0.173)	-1.228*(0.152)	-1.469*(0.283)
大学以上	-1.640*	(0.196)	-1.744*	(0.181)	-1.605*(0.152)	-1.419*(0.238)
その他・無回答	0.445*	(0.216)	0.625*	(0.208)	0.369 (0.214)	0.690*(0.259)
就業状況 [基準=就業経験なし]						
休職中・現在無職	0.169	(0.197)	-0.098	(0.184)	0.019 (0.205)	-0.260 (0.266)
有職	0.089	(0.209)	-0.171	(0.193)	-0.161 (0.212)	-0.570*(0.282)
無回答	0.910	(1.032)	0.651	(1.034)		
父親 [基準=健在]						
死亡			1.197*	(0.148)	0.359*(0.105)	0.772*(0.194)
無回答			6.652*	(0.736)	2.137*(0.280)	2.699*(0.330)
母親 [基準=健在]						
死亡	1.366*	(0.125)				
無回答	6.555*	(1.026)				
標本数	6985		6985		6981	6981
うち無効回答あり	897		1044		965	354
-2 log L	4187.526		4687.597		5279.375	2551.501
$\chi^2$	1168.130*		1204.624*		329.742*	249.661
DF	21		21		20	20

係数の推定値 (標準誤差) を示す。\*:  $p < 0.05$ .  
 父学歴・父職業に関しては本人就業状況が無回答の4ケースをのぞいた。

(a) 母・父の属性

無効回答率の高かった父母の出生年, 父の学歴・職業からとりあげる。表 16 はこれらの項目が無効回答になる確率をロジスティック回帰分析で予測したものである。表 8 と同様の変数を投入しているが、さらにこの項目の回答に大きく影響していそうな「母/父が健在かどうか」を分析に加えることにした。

母出生年の無効回答には、母が健在であるかどうか大きな影響をあたえている。母がすでに死亡している人は、母出生年を答えてくれない確率が高いのである。そしてこれ

をコントロールしたうえでも本人年齢・未既婚・学歴が有意な効果を持っている。高年齢の者、配偶者と離死別したもの、低学歴の者、(有意ではないが)ひとり暮らしの者に無効回答が多い。逆に未婚の者は無効回答がすくなくなっている。

父出生年についても同様のことがいえる。すなわち、父の死亡というファクターの影響が大きく、これをコントロールしたうえでも、本人の年齢と未既婚・学歴・居住人数が有意な効果を持っている。

これらに対して父学歴、父職業の場合はすこし結果がちがう。学歴やひとり暮らしの効果は父出生年の場合とだいたいひとしい。父が死亡していると無効回答が増えることも父出生年の場合と同様である(ただし効果は小さい)。ところが年齢の効果は父出生年とは逆であり、年齢が上がると無効回答が減る傾向にある。また、配偶者との離死別が無効回答を増やすのは父出生年と共通だけれども、未婚であることが無効回答を減らす効果はみられない。

本人の年齢・学歴と父親の学歴・職業との関連についてくわしく検討しておこう。

表-17 本人学歴・年齢と父学歴無効回答率

年齢	中学	高校	専門	短大	大学	その他	無回答
28-37	31.0 (42)	15.3 (594)	14.3 (126)	11.3 (311)	6.2 (321)	25.0 (8)	37.5 (8)
38-47	21.7 (115)	11.6 (665)	13.0 (100)	8.7 (231)	5.3 (323)	8.3 (12)	57.1 (7)
48-57	23.6 (390)	13.8 (788)	11.7 (94)	9.4 (127)	7.0 (229)	23.1 (13)	29.6 (27)
58-67	24.0 (550)	9.1 (529)	2.2 (46)	5.1 (59)	3.6 (138)	25.0 (4)	14.8 (27)
68-77	20.7 (593)	11.4 (334)	14.3 (7)	0.0 (63)	4.0 (50)	0.0 (4)	34.0 (50)
年齢計	22.8 (1690)	12.5 (2910)	11.8 (373)	8.8 (791)	5.7 (1061)	17.1 (41)	30.3 (119)

% (人数) を示す。

父学歴については、表17から、本人学歴の効果がきわめて強く出ていることがわかる。年齢の効果は一貫したものではないが、中学・短大卒の層では若年層で父学歴が無効回答になりやすい傾向がみられる。

表-18 本人学歴・年齢と父職業無効回答率

年齢	中学	高校	専門	短大	大学	その他	無回答
28-37	16.7 (42)	3.9 (594)	4.0 (126)	2.3 (311)	1.9 (321)	25.0 (8)	12.5 (8)
38-47	13.9 (115)	2.6 (665)	1.0 (100)	0.9 (231)	1.5 (323)	8.3 (12)	14.3 (7)
48-57	7.4 (390)	3.4 (788)	3.2 (94)	3.1 (127)	1.7 (229)	0.0 (13)	11.1 (27)
58-67	10.9 (550)	4.3 (529)	4.3 (46)	3.4 (59)	2.9 (138)	0.0 (4)	22.2 (27)
68-77	10.3 (593)	5.1 (334)	0.0 (7)	1.6 (63)	6.0 (50)	0.0 (4)	22.0 (50)
年齢計	10.2 (1690)	3.7 (2910)	2.9 (373)	2.0 (791)	2.1 (1061)	7.3 (41)	18.5 (119)

% (人数) を示す。

父職業(表18)に関しては、中学校卒業者で無効回答率が1割と高めになっているが、それ以外の学歴層は1-4%程度でほとんど差がない。各学歴層内部でみると、中学校卒業

者では若年齢層で無効回答が多く、大学卒業者や学歴無回答の場合は逆に高年齢層で無効回答が多くなる傾向がみられる。

(b) 本人の属性

表 19 は本人の属性についてのロジスティック回帰分析結果である。問 22 の「この 1 年間の健康状態」は属性とはいえないかもしれないが、便宜上ここに置いてある。

表-19 年収・学歴・離家経験・健康の無効回答を予測するロジスティック回帰

	世帯年収 (Q15)	本人年収 (Q10)	学歴 (Q0401)	離家経験 (Q07)	健康 (Q22)
(定数項)	-3.139*(0.346)	-3.562*(0.658)	-5.786*(0.762)	-5.167*(0.865)	-2.611*(0.593)
性別 [男=1, 女=2]	0.569*(0.097)	0.227 (0.193)	0.082 (0.171)	0.080 (0.232)	-0.404*(0.164)
年齢 [基準=28-37才]					
38-47	-0.432*(0.133)	-0.032 (0.289)	0.066 (0.349)	1.139*(0.581)	0.363 (0.328)
48-57	-0.328*(0.139)	-0.136 (0.295)	0.703*(0.311)	0.984 (0.582)	0.198 (0.329)
58-67	-0.394*(0.159)	-0.091 (0.318)	0.819*(0.334)	2.011*(0.564)	0.871*(0.326)
68-77	-0.141 (0.169)	-0.321 (0.356)	1.859*(0.331)	1.164 (0.603)	0.876*(0.348)
同居人数 (本人ふくむ) [基準=2人]					
1人	-1.946*(0.435)	-0.194 (0.492)	0.747 (0.384)	0.823 (0.454)	-0.468 (0.388)
3人	0.693*(0.144)	0.580*(0.256)	0.070 (0.242)	0.290 (0.307)	0.026 (0.202)
4人	0.878*(0.149)	0.143 (0.297)	0.128 (0.265)	0.013 (0.383)	-0.356 (0.257)
5人	0.978*(0.161)	0.070 (0.343)	0.110 (0.297)	0.507 (0.374)	-0.265 (0.282)
6人以上	1.641*(0.143)	0.685*(0.279)	0.100 (0.267)	0.548 (0.327)	-0.291 (0.248)
未既婚 [基準=有配偶]					
離死別	0.524*(0.144)	0.271 (0.289)	-0.209 (0.296)	0.234 (0.332)	0.128 (0.248)
未婚	1.252*(0.143)	0.123 (0.346)	-0.311 (0.406)	0.019 (0.537)	0.456 (0.322)
学歴 [基準=中学校]					
高校・専門	-0.376*(0.114)	-0.461*(0.223)		0.039 (0.247)	-0.510*(0.179)
短大・高専	-0.461*(0.164)	-0.739*(0.370)		-1.004 (0.623)	-1.423*(0.440)
大学以上	-0.510*(0.168)	-0.356 (0.316)		-0.994 (0.554)	-1.095*(0.313)
その他・無回答	0.695*(0.263)	1.433*(0.356)		1.967*(0.363)	1.187*(0.297)
就業状況 [基準=就業経験なし]					
休職・無職	-0.142 (0.217)	-0.433 (0.394)	0.633 (0.604)	-0.516 (0.403)	-0.314 (0.354)
有職	-0.599*(0.226)	-0.771 (0.413)	1.270*(0.614)	-1.023*(0.437)	-0.478 (0.374)
標本数	6981	6981	6981	6981	6981
うち無効回答あり	678	142	159	97	190
-2 log L	4029.123	1341.782	1457.485	920.923	1625.952
$\chi^2$	420.708*	45.524*	59.556*	101.313*	118.320*
DF	18	18	14	18	18

係数の推定値 (標準誤差) を示す。\*:  $p < 0.05$ 。本人就業状況が無回答の 4 ケースをのぞいた。

世帯年収に関しては、世帯の規模が大きくなると無効回答が増えるという強い関係がある。配偶者がいない場合にも、無効回答が増える。また学歴が高い場合、有職である場合に無効回答は生じにくい傾向がある。年齢も世帯年収の無効回答に影響しているが、線形ではなく U 字型の関連 (若年者層と高齢者層で無効回答が多く、中間の年齢で無効回答がすくない) である。

学歴・年齢と世帯年収の無効回答率を検討してみると (表 20)、交互作用が認められる。中学および「無回答」では年齢があがるにしたがって無回答が増加するのに対して、それ

ら以外のカテゴリーでは逆に年齢があがると無回答が減る傾向にある。

表 21 は世帯類型と世帯年収の無効回答率との関連である。ひとり暮らしの場合に無効回答がすくないことがわかる。一方、無効回答が多いのは、高齢の離死別者層と若年の未婚者層である。

表-20 学歴・年齢と世帯年収無効回答率

年齢	中学	高校	専門	短大	大学	その他	無回答
28-37	9.5 (42)	13.8 (594)	15.9 (126)	15.8 (311)	9.7 (321)	12.5 (8)	12.5 (8)
38-47	9.6 (115)	8.1 (665)	8.0 (100)	9.1 (231)	6.2 (323)	8.3 (12)	0.0 (7)
48-57	10.3 (390)	6.6 (788)	12.8 (94)	5.5 (127)	3.5 (229)	7.7 (13)	18.5 (27)
58-67	11.3 (550)	6.4 (529)	2.2 (46)	5.1 (59)	2.9 (138)	0.0 (4)	11.1 (27)
68-77	15.7 (593)	9.9 (334)	0.0 (7)	3.2 (63)	6.0 (50)	0.0 (4)	26.0 (50)
年齢計	12.4 (1690)	8.8 (2910)	11.0 (373)	10.4 (791)	6.2 (1061)	7.3 (41)	18.5 (119)

% (人数) を示す。

表-21 未既婚・世帯人数・年齢と世帯年収無効回答率

年齢	有配偶		離死別			未婚	
	単独	複数	単独	2人	3人以上	単独	複数
28-37	0.0 (3)	10.4 (968)	0.0 (12)	10.0 (10)	20.7 (29)	0.0 (63)	24.6 (325)
38-47	0.0 (1)	7.7 (1252)	0.0 (10)	10.0 (20)	2.7 (37)	0.0 (29)	14.4 (104)
48-57	0.0 (6)	7.7 (1473)	0.0 (29)	6.7 (45)	5.7 (53)	4.0 (25)	10.8 (37)
58-67	0.0 (4)	7.4 (1159)	0.0 (49)	12.2 (49)	21.5 (65)	0.0 (17)	10.0 (10)
68-77	0.0 (1)	11.7 (810)	4.8 (83)	15.1 (53)	27.3 (132)	5.3 (19)	0.0 (3)
年齢計	0.0 (15)	8.7 (5662)	2.2 (183)	11.3 (177)	19.0 (316)	1.3 (153)	20.9 (479)

% (人数) を示す。

表-22 学歴・年齢と本人年収無効回答率

年齢	中学	高校	専門	短大	大学	その他	無回答
28-37	2.4 (42)	1.3 (594)	3.2 (126)	1.9 (311)	1.9 (321)	0.0 (8)	12.5 (8)
38-47	2.6 (115)	1.7 (665)	0.0 (100)	1.7 (231)	2.5 (323)	0.0 (12)	0.0 (7)
48-57	2.1 (390)	2.0 (788)	2.1 (94)	0.8 (127)	1.3 (229)	0.0 (13)	3.7 (27)
58-67	3.3 (550)	2.1 (529)	0.0 (46)	0.0 (59)	0.7 (138)	0.0 (4)	7.4 (27)
68-77	2.9 (593)	1.2 (334)	0.0 (7)	0.0 (63)	0.0 (50)	0.0 (4)	14.0 (50)
年齢計	2.8 (1690)	1.7 (2910)	1.6 (373)	1.4 (791)	1.7 (1061)	0.0 (41)	9.2 (119)

% (人数) を示す。

本人年収の無効回答の出かたは、世帯年収とは大きくちがっている（表 19 左 2 列目）。まず年齢の効果がみられない。同居人数の効果はあるが、3人の世帯か6人以上の世帯の場合に無効回答が多いという結果で、解釈が困難である。学歴は、短大・高専で無効回答がいちばんすくなるという非線形の関係である。ただし有職の場合に無効回答が出にくいことは世帯年収と共通である。

学歴・年齢と本人年収の無効回答率との関連も検討してみたが(表 22)、学歴が「無回答」の場合に本人年収の無効回答が多いという以外には、さしたる関連はなかった。

表 19 にもどって、本人学歴の無効回答については、年齢があがるほど無効回答率があがるという線形の関連がある。また、ひとり暮らしの者は学歴の無効回答率が高い。さらに、有職者の場合に無効回答になる傾向が出ている。

就業状況・年齢と学歴の無効回答率との関連をみると(表 23)、有職かあるいは休職中の場合に高齢者層で無効回答率が高まることがわかる。

表-23 就業状況(Q08)・年齢(Q02S)と本人学歴無効回答率(Q0401)

年齢	有職	休職	現在無職	未経験	無回答
28-37	1.4 (1083)	0.0 (18)	0.3 (303)	0.0 (6)	— (0)
38-47	1.4 (1220)	0.0 (18)	1.0 (201)	0.0 (14)	— (0)
48-57	2.7 (1353)	5.3 (19)	0.7 (272)	0.0 (24)	— (0)
58-67	2.6 (703)	6.3 (16)	1.9 (565)	0.0 (67)	50.0 (2)
68-77	7.6 (236)	17.6 (17)	4.0 (757)	3.4 (89)	0.0 (2)
年齢計	2.3 (4595)	5.7 (88)	2.2 (2098)	1.5 (200)	25.0 (4)

% (人数) を示す。

離家経験については、単身者で無効回答が多いほか、高学歴者・有職者で有効回答が多くなる傾向がみられる(表 19 右から 2 列目)。

健康については、高齢者層で無効回答が多い一方で、高学歴者・有職者で有効回答が多くなる傾向がみられる(表 19 右端列)。

### (c) 親族構成・意見・悩み・心身状態

表 12 で検討したとおり、親族構成(問 31)、意見(問 20)、悩み(問 21)、心身状態(問 23)の変数群はそれぞれの内部の相関が高い。それだけでなく、実際の分析では因子得点を求めたり、論理的に合成変数をつくったりしてリスト・ワイズで処理されることが多いだろう。そこで以下の分析ではそれぞれひとまとめにして分析することにしよう。

表 24 に各項目群のなかでの有効回答の個数の度数分布を示す。ひとつでも欠損変数があるケースは除外するという listwise の原則に沿って分析したとすると、意見や悩みでは 4%、親族構成や心身状態では 10% のケースが分析からぬけおちることになる。

表 25 はこれらの 4 つの項目群についての無効回答の発生を予測するロジスティック回帰分析の結果である。目的変数となっているのは、項目群を構成する項目のなかにもひとつでも無効回答があるか/すべて有効かのロジットである。

親族構成については、女性にくらべて男性のほうが無回答になりやすい傾向がある。若年層・高学歴層では有効回答になりやすい。

意見の項目群では高齢者・中学校卒業者・職業未経験者が無回答になりやすい。

表-24 親族構成・意見・悩み・心身状態の有効回答項目数の度数分布

有効項目数	親族構成 (問 31A-C, E-G)	意見 (問 20)	悩み (問 21(工)(才))	心身状態 (問 23)
0	29 (0.4)	43 (0.6)	170 (2.4)	43 (0.6)
1	44 (0.6)	47 (0.7)	92 (1.3)	18 (0.3)
2	62 (0.9)	12 (0.2)	6723 (96.2)	3 (0.0)
3	81 (1.2)	11 (0.2)		10 (0.1)
4	99 (1.4)	46 (0.7)		8 (0.1)
5	356 (5.1)	140 (2.0)		7 (0.1)
6	6314 (90.4)	6686 (95.7)		6 (0.1)
7				5 (0.1)
8				4 (0.1)
9				5 (0.1)
10				9 (0.1)
11				9 (0.1)
12				16 (0.2)
13				30 (0.4)
14				107 (1.5)
15				451 (6.5)
16				6254 (89.5)

% (人数) を示す。

悩みの項目群では、意見の項目群と同様に、高齢者・中学校卒業者・職業未経験者で無回答になりやすい傾向がある。それらに加えて単身者・離死別者・未婚者が無回答になる傾向が強く出ている。これらのなかでも特に高齢者・単身者・未婚者の無回答傾向は非常に顕著である。

心身状態の項目群でも、高齢者・中学校卒業者で無回答になりやすい傾向がある。

## 5. まとめと提言

### (1) 無効回答要因

分析からあきらかになったことをまとめておこう。

まず、学歴の効果が一貫して高かったことが重要である。分析対象とした50個の項目すべてについて学歴の低い層で無効回答が多い傾向があった。NFR98データの義務教育修了者の回収率が低いことはすでに指摘されている(日本家族社会学会, 2000: 16)が、それに加えて無効回答のために分析から脱落してしまうことが多いことになる。NFR98データの分析結果は、高学歴層を過剰に代表したものになるおそれ大きい。

年齢も重要な要因といえる。ほとんどの項目で、高齢者層の無効回答が多くなる傾向が確認できた。ただし年齢の効果はすべての項目について一貫していたわけではなく、収入および父の属性に関する項目では年齢によるちがいはみられなかった(表19, 表16)。

これらのほか、同居人数・未既婚・就業状況・性別も、項目によっては無効回答をもたらす要因となっていた。ただしこれらの要因の効果は項目によってちがっていた。たとえば単身者が無効になりやすい項目もあれば、有効になりやすい項目もあった。

表-25 親族構成・意見・悩み・心身状態の無効回答を予測するロジスティック回帰

独立変数	親族構成 (問 31)	意見 (問 20)	悩み (問 21)	心身状態 (問 23)
(定数項)	-1.840*(0.343)	-3.032*(0.477)	-4.677*(0.541)	-2.738*(0.329)
性別 [男=1, 女=2]	-0.299*(0.091)	-0.095 (0.133)	-0.234 (0.151)	0.156 (0.088)
年齢 [基準=28-37才]				
38-47	0.447*(0.169)	0.648*(0.258)	0.980*(0.342)	0.187 (0.156)
48-57	0.584*(0.164)	0.590*(0.261)	1.723*(0.327)	0.575*(0.148)
58-67	0.703*(0.175)	1.021*(0.268)	2.253*(0.339)	0.783*(0.158)
68-77	0.652*(0.188)	1.018*(0.285)	2.257*(0.353)	0.741*(0.172)
同居人数 (本人ふくむ) [基準=2人]				
1人	0.198 (0.200)	0.167 (0.295)	1.250*(0.228)	-0.079 (0.207)
3人	-0.041 (0.124)	0.313 (0.177)	0.108 (0.212)	-0.018 (0.117)
4人	-0.021 (0.135)	0.183 (0.202)	0.066 (0.251)	0.081 (0.126)
5人	-0.095 (0.155)	0.228 (0.223)	0.261 (0.264)	0.065 (0.144)
6人以上	0.229 (0.134)	0.371 (0.194)	0.314 (0.237)	0.034 (0.134)
未既婚 [基準=有配偶]				
離死別	0.444*(0.135)	0.134 (0.201)	0.933*(0.195)	0.116 (0.136)
未婚	-0.080 (0.197)	0.375 (0.263)	2.288*(0.235)	0.125 (0.178)
学歴 [基準=中学校]				
高校・専門学校	-0.452*(0.100)	-0.691*(0.150)	-0.362*(0.164)	-0.224*(0.099)
短大・高専	-0.811*(0.185)	-0.770*(0.267)	-0.691*(0.325)	-0.166 (0.156)
大学以上	-1.378*(0.186)	-0.698*(0.228)	-0.374 (0.261)	-0.452*(0.155)
その他・無回答	1.196*(0.206)	1.389*(0.242)	1.818*(0.267)	1.226*(0.208)
就業状況 [基準=就業経験なし]				
休職中・現在無職	-0.167 (0.224)	-0.432 (0.273)	-0.217 (0.290)	0.114 (0.214)
有職	-0.137 (0.233)	-0.583*(0.288)	-0.785*(0.314)	-0.138 (0.223)
標本数	6981	6981	6981	6981
うち無効回答あり	670	299	262	731
$-2 \log L$	4154.498	2319.057	1771.716	4513.002
$\chi^2$	259.555*	149.951*	462.409*	168.685*
DF	18	18	18	18

係数の推定値 (標準誤差) を示す。\*:  $p < 0.05$ . 本人就業状況が無回答の4ケースをのぞいた。

すでに実施された調査に関しては、出てしまった無効回答はどうしようもない。無効回答に由来するバイアスを適切に評価することを考えるしかない。上記のような、無効回答の発生と密接にかかわっている変数を投入した分析結果を解釈する場合、欠損データによるゆがみが影響している可能性がないかどうかをチェックしてみることをおすすめする。

## (2) 無効回答を減らすために

今後同様の調査をおこなう場合、調査法上の工夫で無効回答を減らせるだろうか。今回の分析から、効果がありそうな無効回答対策を考えてみよう。

### (a) 調査票回収時のチェック

NFR98 調査は、対象者自身が調査票を読んで回答を記入する自記式の調査であった。調査員が対象者宅を訪問して、調査の趣旨を説明して調査票を手渡し、対象者が記入したものをあとで調査員が回収する (その際にチェックをおこなう) という方式である。個々の質問について調査員がたずねて記入する他記式の調査にくらべると回答者の負担が大き

いので、たとえば調査票の一部を飛ばしてしまったり、質問内容を取りちがえたりといった間違いが起こりやすい。このような調査法上の特性に考慮して、重要な項目やまちがいやすい項目については調査員が回収時にチェックすることになっている。

表11でみたように、調査票回収時にこのチェックが入った項目は無効回答がすくない。そこで無効回答を減らすために回収時チェックを意識的に使うことを提言しておきたい。先行研究やプリ・テストで無効回答の発生率が高いことが予測できる項目を洗い出して、それらをチェック項目に加えておけばよい。

#### (b) 回答選択肢の再検討

表11からわかるように、特に無効回答の多かった項目は、回答欄に数値を直接書きこむ項目と、回答選択肢に「その他」「わからない」をふくむ項目とで占められている。これらの回答方式を避けることによって、無効回答を減らすことができるだろう。

- 数量を答えてもらうときは数値をそのまま書くのではなくて、適当な区間を区切った選択肢から選んでもらうようにする。
- 回答選択肢にはなるべく「その他」「わからない」といった選択肢を設けないようにする。もちろん網羅的な選択肢を用意できないとか、他人や世帯に関する事項で回答者が本当に知らないケースがあるから、こうした選択肢を完全に除外することはできない。そうした項目については、回収時の調査員チェックを必ず入れて、できるだけこまかい情報を聞き出す努力が必要である。

#### (c) 質問文作成上の工夫

本稿の分析では学歴の低い人の無効回答が一貫して多く、また多くの項目で高齢者の無効回答が多くなっていた。質問文を読んで該当する選択肢を選ぶ、というような解答形式は学校での試験問題の形式と共通しているため、学校文化に慣れ親しんだ人ほど答えやすいのかもしれない。また高齢者の無効回答が多い原因は、視力が低下しているなどのために質問文を読んで回答することの負担が大きいためかもしれない<sup>2)</sup>。

これらの推測があたっているとすれば、回答選択肢の数を減らしたり質問文を短くしたりする、文字を大きくする、同じ形式の連続する質問では行間を空ける、などのように、回答者の負担を減らす工夫によって無効回答を減らす効果が期待できる。

#### (d) 地域別の対応

表3で検討したように、調査地域によって無効回答率がかなりちがう。また図1では、回収率の高い地域で無効回答率が高くなっている場合のあることもわかった。なぜこのような差が出るのかについては今後の研究をまちたい。ここでは仮説的にいえそうなことを書いておこう。

図1にあらわれた回収率と無効回答率の相関は、つぎのように解釈できる。回収率の低い地域では調査に協力的な層と非協力的な層がはっきりとわかれていて、非協力的な層が調査そのものを拒んでいるのではないだろうか。これに対して回収率の高い地域は、かならずしも調査に積極的に協力する人たちが多いわけではなく、調査そのものには応じるけれども回答の動機づけの弱い、いわば消極的な調査協力者が多いのかもしれない。

そうとすると、地域によってちがう対応をとる必要がある。回収率が高くても無効回答の多い地域（図1の山形・鳥取・徳島など）では、むやみに回収率をひきあげるよりは、無効回答をへらして精度の高い回収票をえることが重要になってくる。調査員は、いちおう調査に協力してもらえたケースについて、回答者の動機づけをできるだけ高める、あるいは答えにくい質問についてしっかりと確認をおこなうといった対応を心がけなければならない。無効回答の多い地域には人件費を重点的に配分して優秀な調査員を確保するという対策も有効かもしれない。

## 注

- (1) 保田時男（大阪大学大学院人間科学研究科）氏の示唆による。
- (2) 休業者の無効回答が多い（表4）ことについても、高齢者と同様の解釈ができる。

## 文献

- 日本家族社会学会 全国家族調査研究会、1999 『1998年「家族についての全国調査」(NFR98)コード・ブック』。
- 日本家族社会学会 全国家族調査(NFR)研究会、2000 『家族生活についての全国調査(NFR98) No. 1』。
- SAS Institute、1990 『SAS Technical Report P-200: SAS/STAT Software: CALIS and LOGISTIC Procedures, Release 6.04』、ISBN 1-55544-390-7。
- 保田 時男 2000 「クロス集計表における欠損データの分析：学歴移動表を例として」『理論と方法』15: 165-179、ISSN 0913-1442。