

コミュニティ福祉学部生を対象とした質問紙調査データの一般化可能性について¹

— GHQ 精神健康調査票への回答傾向からの検討と考察 —

嘉瀬 貴祥²

(コミュニティ福祉学研究科／日本学術振興会特別研究員 DC2)

1. 諸言

質問紙調査は、調査研究における一般的な手法となっている。特に、大学生を対象とした質問紙調査はその容易さもあり、心理学研究において実施されることが多いように見受けられる。ところで、質問紙調査の対象者の選定、すなわちサンプリングの方法については様々な議論がなされている (e.g., 土屋2005)。サンプリングは「どのような母集団を想定した研究を行うか」という重要な問題と関連している。例えば、立教大学コミュニティ福祉学部 に所属する大学生のみを対象とした調査の結果を用いて行った卒業研究の題目を「コミュニティ福祉学部 に所属する大学生の精神的健康について」とする場合は少なく、たいていの場合は「大学生における精神的健康について」と一般的な対象を想定して論を進めることが多い。このように大きな母集団を想定して研究を進める場合には、単純無作為抽出や二段抽出など、想定する母集団からできるだけ無作為に対象者を選定するような方法で、サンプリングを行うことが求められる (南風原2002)。しかしながら、全国の大学生を対象として調査を実施することは困難であり、先述した卒業研究の例の様に、自分が所属する大学や周辺の大学に所属する大学生へ依頼することもあるだろう。このように限定的な調査対象者より得られたデータの場合、調査対象者に共通する属性が回答の傾向に反映されることがある。仮に一般的でない回答の傾向があった場合、データの分析結果はこれに左右されるため、そこから得られた知見を大きな母集団を想定した研究の結果として断定することは適切でないと考えられる。現実的には、研究結果の一般化可能な範囲 (一般化可能性) を調査対象者の属性や先行研究との比較から判断し、得られた分析結果を適切に考察することが求められる。つまり土屋 (2005) が指摘するように、調

査を実施した集団特性を考慮した考察が必要となってくるということである。

これらの事項をふまえると、一大学の一学部、例えば立教大学コミュニティ福祉学部において得られたデータは、どれほどの範囲で一般化が可能であるかという関心が見出される。そこで本研究では、①コミュニティ福祉学部にも所属する大学生を対象とした質問紙調査（以下、調査1（コミ福）と略記）の結果を、②立教大学に設置されている10学部（文学部、異文化コミュニケーション学部、経済学部、経営学部、理学部、社会学部、法学部、観光学部、コミュニティ福祉学部、現代心理学部）にも所属する大学生を対象とした調査（以下、調査2（立教）と略記）、③全国の大学生を対象としたインターネット調査（以下、調査3（Web）と略記）の結果、すなわち標本抽出の無作為性が異なる3件の調査の結果を比較することで、その一般化可能性について検討することとした。これら3件の調査結果を、以下の観点より比較検討する。第一に質問項目に対する回答の分布の観点、第二に尺度得点の分布の観点である。測定尺度には、日本語版GHQ精神健康調査票12項目版（中川・大坊2013；以下、GHQ12と略記）を用いることとした。GHQ12は高い信頼性と妥当性が示されているうえ、その適応可能性について、健常者、神経症者、大学生などの多様な母集団から得たデータをもとに検討されている点で、本研究の目的と合致している。本研究では、上述した3件の調査結果に加え、中川他（2013）が報告しているGHQ12を用いた大規模調査の結果との比較を行うことで、コミュニティ福祉学部の学生を対象とした質問紙調査の一般化可能性について考察することとした。

2. 方法

1) 調査内容

本研究では、調査対象者の異なる3件の質問紙調査の結果を比較することとした。

①調査1（コミ福）

立教大学コミュニティ福祉学部にも所属する大学生268名を対象として、2015年10月に集合調査法による質問紙調査を行った。この調査対象者から200名（男性50名、女性150名、平均年齢19.7歳、 $SD=1.3$ ）を無作為に抽出し、分析対象とした。なおこの調査は、立教大学コミュニティ福祉学部倫理委員会の承認（番号：2015-005）を得て実施された。

②調査2（立教）

立教大学に設置されている10学部（文学部、異文化コミュニケーション学部、経済学部、経営学部、理学部、社会学部、法学部、観光学部、コミュニティ福祉学部、現代心理学部）にも所属する大学生561名を対象として、2013年10月に集合調査法による質問紙調査を行った。この調査対象者から200名（男性85名、女性

115名、平均年齢19.9歳、 $SD=1.2$)を無作為に抽出し、分析対象とした。なおこの調査は、立教大学コミュニティ福祉学部倫理委員会の承認(番号:2013-005)を得て実施された。

③調査3 (Web)

株式会社クロス・マーケティングの調査モニターである全国の大学生400名を対象として、2016年6月にインターネット調査を行った。この調査対象者から200名(男性102名、女性98名、平均年齢19.6歳、 $SD=1.1$)を無作為に抽出し、分析対象とした。なおこの調査は、立教大学コミュニティ福祉学部倫理委員会の承認(番号:2016-2)を得て実施された。

2) 測定尺度

諸言にて述べたように、GHQ12(中川他2013)を用いた。この尺度は、個人の精神的健康を測定する尺度であり、「いつもより日常生活を楽しく送ることが」など12項目の質問からなり、「できた」、「いつもと変わらなかった」、「できなかった」、「全くできなかった」などの4段階で自己評価を行うものである。採点には、「できた」と「いつもと変わらなかった」を0点、「できなかった」と「全くできなかった」を1点とするGHQ採点法を用いた。GHQ12の得点が高いほど、精神的健康が低いことを示す。なお、GHQ12得点に設定された診断基準によると、4点以上を示すものは精神的健康が不良であると見込まれる(中川他2013)。

3) 統計処理

本研究では、統計解析プログラムHAD15.011(清水2016)およびR3.2.3を用いて分析を行った。すべての分析における有意水準は5%に設定した。先んじて、各調査で得られたGHQ12得点における性差と学年差を確認するため、男女間および各学年間でGHQ12得点についてWelchの t 検定を行った。その結果、どの調査においても男女間および学年間に有意な差は認められなかった。本研究で用いた3件の調査では、調査対象者の男女比や学年比が異なっている。しかしながら、上述したWelchの t 検定の結果を受けて、本研究では性別や学年による偏りは生じていないものと便宜的に扱うこととした。

3. 結果

1) 各調査におけるGHQ12各項目に対する回答傾向の比較

第一に、GHQ12を構成する12項目の質問に対する回答傾向を比較するため、各項目に対する回答の平均値、標準偏差、歪度、尖度、項目反応理論(豊田2002)を用いた識別力および困難度の算出を行った(表1)。なおここでは、調査協力者の回答傾向をより詳細に把握するため、上述したGHQ採点法によって換算された値ではなく素点を用いた。

平均値と標準偏差

3件の調査における平均値の範囲は1.9—3.1、標準偏差の範囲は0.7—0.9であった。平均値に注目すると、その大きさが調査3（Web） \leq 調査2（立教） \leq 調査1（コミ福）であった項目が12項目中7項目であった。また標準偏差については、その大きさが調査1（コミ福） \leq 調査2（立教） \leq 調査3（Web）であった項目が、12項目中9項目であった。すなわち、GHQ12の各項目に対する回答傾向を概観すると、調査1（コミ福）では比較的高い値に回答が集中する傾向にあったと考えられる。

歪度と尖度

全体に、歪度は正、尖度は負の値を示す傾向にあった。歪度に注目すると、項目5、項目9、項目10は調査3（Web）のみ正の値を示していた。また尖度については、項目1では調査1（コミ福）のみ、項目4はすべての調査、項目11では調査3（Web）のみが正の値を示していた。すなわち、全体としての右に歪んだ比較的近だらかな得点の分布を示しているが、調査3（Web）は調査1（コミ福）と比較すると、低い値に回答が集中する傾向がいくつかの項目で認められたと考えられる。

識別力と困難度

Roznowski（1989）と豊田（2002）が示す基準では、識別力の値が0.5未満、あるいは困難度の値が4.0以上を示す場合に、その項目は回答者の能力を適切に測定できていないと判断する。識別力に注目すると、調査1（コミ福）の項目8が、基準を下回る値を示していた。加えて、基準値未満ではないものの、調査1（コミ福）の項目3、項目12の識別力は0.5であった。また困難度については、基準値を上回る値を示した項目はなかったが、調査1（コミ福）の項目5と項目8においては識別力が-1.0以下を示していた。この結果から、調査1（コミ福）においてGHQ12のいくつかの項目は、調査協力者の精神的健康の高低を適切に測定できていなかったと考えることができる。

全体の回答傾向

調査1（コミ福）は調査2（立教）および調査3（Web）と比較するとやや高い値に回答する傾向があるが、なかには精神的健康の低くない者でも高い値に回答をしている項目がいくつかある可能性が示された。

表1 各調査におけるGHQ12の項目分析

項目		平均値 (標準偏差)	歪度	尖度	識別力	困難度
1	コミ福	2.1 (0.7)	0.4	0.3	0.8	1.1
	立教	2.1 (0.7)	0.3	-0.2	0.7	1.1
	Web	2.2 (0.8)	0.3	-0.3	1.6	0.6
2	コミ福	2.0 (0.9)	0.6	-0.4	0.6	1.3
	立教	2.2 (0.9)	0.3	-0.8	0.6	0.7
	Web	2.1 (0.8)	0.4	-0.1	1.6	0.7
3	コミ福	1.9 (0.8)	0.5	-0.4	0.5	1.9
	立教	2.1 (0.9)	0.3	-0.7	0.7	0.8
	Web	2.2 (0.9)	0.4	-0.5	1.9	0.5
4	コミ福	2.2 (0.7)	0.3	0.4	1.0	0.9
	立教	2.1 (0.7)	0.4	0.3	0.9	1.1
	Web	2.1 (0.7)	0.6	0.6	1.7	0.8
5	コミ福	3.1 (0.7)	-0.3	-0.5	1.3	-1.0
	立教	2.8 (0.8)	-0.2	-0.6	1.4	-0.5
	Web	2.1 (0.9)	0.4	-0.7	0.9	0.7
6	コミ福	2.7 (0.8)	-0.1	-0.4	1.2	-0.3
	立教	2.7 (0.8)	-0.2	-0.4	0.8	-0.4
	Web	2.5 (0.9)	-0.1	-0.6	0.9	-0.1
7	コミ福	1.8 (0.7)	0.5	-0.2	1.0	1.5
	立教	1.9 (0.8)	0.6	-0.3	1.5	-0.5
	Web	2.4 (0.8)	0.2	-0.4	1.0	0.3
8	コミ福	2.7 (0.7)	-0.0	-0.2	0.1	-5.3
	立教	1.9 (0.7)	0.4	-0.1	1.0	1.2
	Web	2.5 (0.9)	0.0	-0.8	1.0	0.0
9	コミ福	2.8 (0.8)	-0.2	-0.4	1.6	-0.5
	立教	2.7 (0.8)	-0.2	-0.5	1.6	-0.4
	Web	2.4 (0.9)	0.1	-0.7	1.5	0.2
10	コミ福	2.8 (0.8)	-0.3	-0.3	1.3	-0.5
	立教	2.7 (0.8)	-0.3	-0.3	1.0	1.2
	Web	2.3 (0.9)	0.2	-0.7	1.3	0.3
11	コミ福	2.5 (0.9)	0.1	-0.7	1.0	0.1
	立教	2.4 (0.9)	0.0	-0.7	0.7	0.1
	Web	2.2 (0.8)	0.7	0.6	1.3	0.8
12	コミ福	2.2 (0.7)	0.3	-0.1	0.5	1.1
	立教	2.2 (0.8)	0.3	-0.3	0.6	0.9
	Web	2.2 (0.8)	0.5	-0.2	1.3	0.6

2) GHQ12合計得点の比較

各調査におけるGHQ12得点を比較するため、正規分布を仮定しないWelchの*t*検定を行った。その結果、調査1（コミ福）のGHQ12平均得点は調査3（Web）と比べて有意に高いことが認められた（表2）。その他の調査間には、差が認められなかった。各調査におけるGHQ12得点の平均値と標準偏差を見ると、調査1（コミ福）は平均値が高く標準偏差の値が小さいのに対して、調査3（Web）は平均値が低く標準偏差の値が大きかった。平均値に有意差は認められなかったものの、調査2（立教）の平均値と標準偏差の値はその他2件の調査の中間を示していた。また、調査1（コミ福）は歪度が0であるうえ、正規性が認められた。これに対し、調査2（立教）と調査3（Web）は歪度が正の値を示し、正規性が認められなかった。すなわち、調査1（コミ福）は平均値である5.2点付近の得点を示すものが集中する傾向があるのに対し、調査2（立教）と調査3（Web）は平均値である4.9点あるいは4.3点より低い得点を示す者の割合が多いという傾向があることが示された。この傾向は、GHQ12の各項目に対する回答から読み取れた全体的な傾向と同様であった。

表2 各調査におけるGHQ12得点の記述統計量と一要因分散分析の結果

	平均値 (標準偏差)	歪度	尖度	JB 検定 ^a χ^2	<i>F</i>	一要因分散分析 多重比較 (Holm 法)
コミ福	5.2 (2.8)	0	-0.8	5		
立教	4.9 (3.1)	0.4	-0.6	7.2*	4.7**	コミ福 < CM**
Web	4.3 (3.6)	0.6	-0.8	16.4**		

* $p < .05$, ** $p < .01$

^aJack-Berra検定を指す。

これらの傾向のより容易に把握を可能にするため、各調査におけるGHQ12得点の分布を図示した（図1）。調査1（コミ福）では4—7点、調査2では2—5点、調査3では0—3点の間に、それぞれの調査対象者の約50%が含まれており、上述した得点の傾向を明確に示している。これらの得点の分布を、GHQ12の妥当性と信頼性の検証のために中川他（2013）が実施した大規模調査の結果と比較した。すると、調査1（コミ福）の結果は精神的健康が不良である「神経症者」の示す、4—5点付近にピークがある比較的なだらかな分布と類似していた。一方で調査3（Web）は、精神的健康に問題の無い、あるいは良好である大学生や成人が含まれる「一般」や「健常者」の示す、1—2点付近にピークのある分布と類似していた。調査2（立教）は「神経症者」と「一般」を合成した様な分布であると推察される。このことから、調査1（コミ福）の得点の分布は一般的な傾向とは

異なっており、比較的高めの得点に多くの人数が含まれるような傾向にあることが示された。

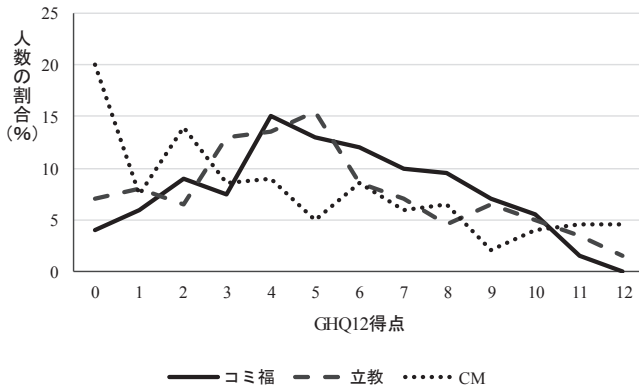


図1 各調査における GHQ12得点の分布の比較

4. 考察

本研究の結果より、3件の調査間でGHQ12に対する回答傾向が異なることが示された。そのなかでもコミュニティ福祉学部 to 所属する大学生は、GHQ12の各項目に対してやや高めに回答する傾向があり、GHQ12得点の分布は中川他（2013）が示す神経症者のGHQ12得点の分布傾向に類似していた。また、立教大学に設置されている複数の学部 to 所属する大学生を対象とした調査の結果は、コミュニティ福祉学部 to 所属する大学生を対象とした調査の結果に比べ、全国の大学生を対象としたインターネット調査や中川他（2013）が報告している一般の調査の結果に類似していた。ただし、コミュニティ福祉学部 to 所属する大学生の得点分布が神経症者の分布と類似していることから、「コミュニティ福祉学部の大学生は全体的に精神的健康に不良である」と即時に判断できるわけではない。先述したように、GHQ12の各項目の識別力と困難度を算出した際に、調査1（コミ福）では調査対象者の精神的健康を正しく測定できていない可能性のある項目がいくつか認められた。つまり、コミュニティ福祉学部 to 所属する大学生はなんらかの要因の影響で、実際の自分の精神的健康の状態より悪い方向に回答をしている可能性も考えられる。これらの知見より、コミュニティ福祉学部 to 所属する大学生のみを対象とした質問紙調査の結果は、少なくとも、全国の大学生まで一般化の範囲を拡大することはできないことが示唆された。これには調査の実施形態（質問紙のデザインや調査時期など）やそれに関連する回答者の動機づけの高低が影響している可能性があるため、一概にコミュニティ福祉学部 to 所属する大学生特有

の傾向であると断定することはできない。しかしながら、ほぼ同様の形態で実施された調査2（立教）とも回答傾向が異なっていたことを考慮すると、所属学部という調査協力者の属性が回答傾向にある程度反映されている可能性は十分に考えられる。

本研究において各調査問に差異が認められたことから、調査結果を集約する際には、回答傾向や得点分布を丁寧に確認し、その得点を用いた分析から算出された結果の解釈は慎重に行う必要があると考えられる。コミュニティ福祉学部注目すると、女子学生の方が男子学生より多い傾向にある。2016年5月時点では、男性790名、女性952名が在籍しており（立教大学2016）、授業やゼミの時間を利用して質問紙調査を実施する場合でも、調査協力者における女子学生の比率が高くなる可能性がある。実際に、2015年に調査を実施した本研究のデータにおいても、女子学生が調査協力者の約7割を占めていた。本研究では性別が回答に大きく影響しないGHQ12を用いたが、例えば先行研究において明確な男女差が認められている心理尺度や、性役割が影響する可能性の高い質問項目を用いて調査を実施する際には、学部における男女比が影響することに十分に留意する必要があると考えられる。しかしながら、このような調査対象者の属性に関する説明は、研究の本来の目的に関連する結果の解釈や考察を複雑化し、研究主旨に対する理解を妨げる場合がある。したがって、大学生の一般的な傾向の調査を目的とする場合には、異なる複数の大学に所属する大学生や、同大学内でも可能な限り複数の学部にも所属する大学生を対象として調査を実施することで、より一般化可能性の高い調査結果を得られるように考慮することも、調査の実施困難度は高くなるものの、重要であることが再確認された。

引用文献

- 南風原朝和（2002）. 心理統計学の基礎 有斐閣アルマ.
- 中川泰彬・大坊郁夫（2013）. 日本語版GHQ精神健康調査票手引き（増補版）日本文化科学社
- Roznowski（1989）. Examination of the measurement properties of the job descriptive index with experimental items. *Journal of Applied Psychology*, 74, pp. 805-814.
- 立教大学（2016）. 各種データ | 立教大学 Retrieved from <https://www.rikkyo.ac.jp/aboutus/profile/data/>（2016年8月8日）
- 清水裕士（2016）. フリーの統計分析ソフトHAD——機能の紹介と統計学習・教育, 研究実践における利用方法の提案—— メディア・情報・コミュニケーション研究, 1, pp. 59-73.
- 豊田秀樹（2002）. 項目反応理論——入門編—— 朝倉書店.
- 土屋隆裕（2005）. 教育心理学研究における調査法 教育心理学年報, 44, pp. 91-99.

脚注

- ¹ 本研究は、日本学術振興会特別研究員(DC2)に採択された研究課題（課題番号：15J01794）の助成を受けて実施された。
- ² 質問紙調査へご協力頂いた皆様、調査の実施にご助力頂いた立教大学の大石和男先生ならびに上村真美先生、統計解析法に関するご助言を頂いた中央大学大学院の飯村周平さんに心より感謝申し上げます。