

視聴質調査における評価尺度の開発 Development of a Scale of Measuring TV viewing and Program Quality—A Web Survey of “Quae”—

○石山 玲子¹，○黄 允一²，岩崎 貞明³，小玉 美意子⁴，戸田 桂大⁵，
中條 浩⁶，藤井 達也⁷，山下 玲子⁸

Reiko ISHIYAMA, Yun'il HWANG, Sadaaki IWASAKI, Miiko KODAMA,
Keita TODA, Hiroshi NAKAJO, Tatsuya HUIII and Reiko YAMASHITA

¹⁴⁵⁸武蔵大学 Musashi University

²フェリス女学院大学 Ferris University, ³メディア総合研究所 Media Research Institute,

⁶株式会社ラビネット Rabinet Co., Ltd, ⁷武蔵大学大学院 Graduate School of Musashi University

要旨…これまで、視聴質調査実施時の評価尺度について開発を進めてきたが、本研究では評価項目の妥当性を検証するために、ネットリサーチ会社に委託し2回にわたる大規模調査を行った。「さまざまな番組を多様な軸で評価する」という調査目的に照らし合わせると、評価項目の妥当性は確保されているといってもよい結果となったものの、因子分析によると若干の修正が必要なことが分かった。そこで、質問項目のクラスター分析を試みることにより、評価項目を再編成するとともに、評価結果によるクラスター分析を行い番組得点を算出した。その結果、番組得点からみる番組評価は、多様なジャンルに適用できる一方、同一ジャンル内での評価も可能なことが明らかになった。

キーワード 視聴質調査, Quae評価尺度, Web調査, 市民参加

1. はじめに

番組の視聴に関する調査は、NHKによる調査を嚆矢として、今日までにビデオリサーチ社をはじめ民間の調査会社やテレビ局、調査団体などが主体となり、量的、質的な調査が古くから実施されてきた。質的な調査（いわゆる視聴質調査）として近年では、リサーチQ、優良放送番組推進会議、テレビウオッチャーなどがインターネットを利用し番組を質的に評価している。しかし、それぞれが独自の目的をもって調査を実施しているものの、それらは視聴者の番組に対する反応を調査するものが主で、番組の特徴を多面的に評価することに重点を置いたものは少ないと思われる。

そこで、武蔵大学 MMS 研究会 Quae プロジェクトでは、2004 年より視聴者の番組に対する反応と番組の質に対する多面的な評価を同時に調査し数量化できる評価尺度の開発、それによる調査と結果の公表を目的として、検討を重ねてきた。それとともに、ネットリサーチ法の開発を行い、2008 年に Quae（クアエ）という名称のインターネットによるテレビの番組評価採点サイトを立ち上げた。2009 年より継続的に運営し（2010 年よりメディア総合研究所との共同研究会）、2か月に1回の定期調査、および特別調査（たとえば、東日本大震災報道やロンドンオリンピックについて）を実施してきた。定期調査は、調査月の末日の夕方から夜にかけて放送される番組を対象とした。また、昨年度は、ジャンルごとに1週間の調査期間を定めた調査も行ってきた。

さらに、2013 年度には新たな取り組みとして、ネットリサーチ会社による2回の委託調査を実施した。目的は、幅広く多くの評価者（およそ 600 人）を得て Quae で開発した指標を用いた評価をしてもらおうと同時に、一般の方々の年代や性別による偏りのないデータを得て、評価尺度の妥当性を再検討する機会を得るためである。これにより指標の確実性を上げ、業界の方々にも使用してもらうなど視聴質調査の浸透を願うものとする。今回はこれらの調査結果を中心に発表を行う。

2. ネットによる視聴質調査Quae（クアエ）の評価項目確立までの取り組み

Quae プロジェクトでは、文献研究・講演・討論などを行う傍ら、2005 年より番組評価項目作成の試みを始めた。ポイントは、質の評価を量的に計測することと、全ての番組に共通する指標づくりにある。まず、尺度項目収集のために、学生を対象に、テレビ番組のイメージについて自由記述式のアンケートを行った。「見たいと思う番組」「見たくないと思う番組」「よい番組」「悪い番組」について、あてはまる条件を5つ以上記述してもらい、その回答をカテゴリー化した。その結果、抽出した60項目を尺度項目とし、実際の番組を視聴後、各項

目を5件法(単極)により評価する質問紙調査を4回実施した。さまざまなジャンルの番組を対象とした予備調査では各番組の特徴に基づく評価がされていること、および、同一ジャンルを複数対象とした予備調査では、同一ジャンルの番組でも評価が可能であることを確認できた。

さらに、予備調査の結果を分析した上で、回答者への負担を軽減する目的で項目数を34項目に絞り込み、2006年に一般サンプルを採用した中規模の質問紙調査を実施した。対象は、関東在住10代~80代の男女374名で、対象番組は朝の情報番組、夜の情報番組、娯楽番組をそれぞれ2番組ずつ計6番組で結果を比較、分析した。その結果、調査項目の有効性を確認することができた。さらに、評価項目をその内容から慰安評価・品質評価・実用評価・倫理評価という4つの軸に振り分け、4軸で番組の特性を評価する形式を採用した。

2008年の大規模インターネット調査に先立ち、2007年には、より簡便な調査の開発を目的として、さらに評価項目を精査し20項目まで削減するとともに、調査方法の検討を行った。評価項目については対象としたどの番組に対しても一定の評価がされており、詳細に分析した結果、項目の有効性が確認された。調査方法では、5件法単極・5件法両極・2件法による違いをみるため、回答者を同質の3つのグループに分け調査を行い検討したところ、5件法単極が評価をする上でふさわしいという結論となった。その後、2010年に評価項目の再検討を重ね若干の修正をしつつ、現在の評価項目(4軸20項目)を使用した調査に至っている。

3. 研究方法

ネットモニターを対象とした2回の委託調査のうち、第1回の調査実施期間は2013年9月26日(木)から28日(土)の3日間で、ドラマ番組を対象に実施した(以後、ドラマ調査とする)。回答者は年代別に6区分し(15歳~24歳、25歳~34歳、35歳~44歳、45歳~54歳、55歳~64歳、65歳以上)、それぞれ104人ずつ、合計624人で男女半々である。調査対象番組の選定は、2013年7月~9月に放送されたドラマの内、地上波およびNHKのBSプレミアムで19時から24時に放送された日本のドラマで45分以上の番組という条件で行ったところ、28番組(NHK6番組、民放18番組、BS4番組)が評価対象となった。

続く第2回では、2013年12月31日の大晦日の17時から24時までの間に30分以上放送された地上波の番組を対象に、2014年1月1日(水)に調査を実施した(以後、大晦日調査とする)。回答者は第1回目と同様624人で、評価対象となった番組は28番組(NHK11番組、民放17番組)であった。

両調査では、まず、作成した番組リストを提示し、その中から、「よく見た番組の順に3番組まで」を選んでもらい、Quaeで開発した評価尺度を用いた20項目に従って、番組を5件法・単極(1点:あてはまらない~5点:あてはまる)で評価してもらった。使用したQuaeの評価項目は、品質評価(Quality)、娯楽評価(Amusement)、実用評価(Usefulness)、倫理評価(Ethics)と4つの軸に分けられ、具体的には、品質評価として「映像や音楽が良い」「演出が良い」「品が良い」「独創性がある」「構成が適切である」という5項目、さらに、娯楽評価として「楽しい」「リラックスする」「感動できる」「共感できる」「出演者が好き」という5項目である。加えて、実用評価としては「見ごたえがある」「世の中のことがわかる」「生活に役立つ」「教養が身につく」「話題性がある」という5項目、それに、倫理評価(5項目)として「公正さにかける」「良識に反する」「暴力表現が過剰である」「性表現が過剰である」「差別的表現がある」(すべて逆転項目)という5項目となっている。

両調査では、そのほか、それぞれの番組の視聴量(全部視聴・ほとんど視聴・半分くらい視聴・少し視聴・1回のみ視聴)について尋ねた。さらに、ドラマ調査では、視聴機器(テレビ・パソコン・携帯端末・その他)、リアルタイム視聴の有無(リアルタイム・生で視聴、録画して視聴、有料動画サイトで視聴、無料動画サイトで視聴、その他)について尋ねるとともに、普段の視聴状況として好みのジャンルについても調査した。

両調査結果を番組ベースで見ると、ドラマ調査では延べ1623件、大晦日調査では1586件の回答が得られた。双方における、番組ベースの回答者の性別はおおよそ男女半々で、年代別にも大きな偏りは見られなかった。

4. 研究結果

2つの調査では、「さまざまな番組を多様な軸で評価する」という調査目的に照らし合わせると、評価項目の妥当性は確保されているといってもよい結果となった。ただし、詳細に分析を行うと、これまでのQuaeでの独自調査の仮定と若干食い違う点がみられることが分かった。たとえば、両調査結果をみると、20の質問項目は概ね3つないし4つの因子に分れることが明らかになったものの、これまでQuaeの4つの評価軸として採用していた項目群とは完全には一致しないことが分かった。たとえば、「品質」としてひとかたまりに扱ってきた項目群のうち「品がよい」という項目は、「娯楽」として扱ってきた「感動できる」や「共感できる」という項目と同じグループとして扱った方がよいというような例や、「娯楽」としてひとかたまりに扱ってきた項目群のうち「出演者が好き」という項目は、「実用」として扱ってきた「話題性がある」や「見ごたえがある」という項目と同じグループとして扱った方がよいというような例が散見された。

そこで、質問項目について類似・非類似度に基づいたまとまりを明らかにするために、2つの調査結果における質問項目のクラスター分析を試みた。これにより、番組評価の基準がどのような観点から行われているか明らかにになり、質問項目をまとめることが可能となる。その結果、両調査に共通して、おおよそ6つのクラスターに

分類できることが分かった。それぞれには、「楽しい」「リラックスできる」「感動できる」「倫理的に問題がある」「実用的である」「作りが良い・見ごたえがある・話題性がある」と命名できる。

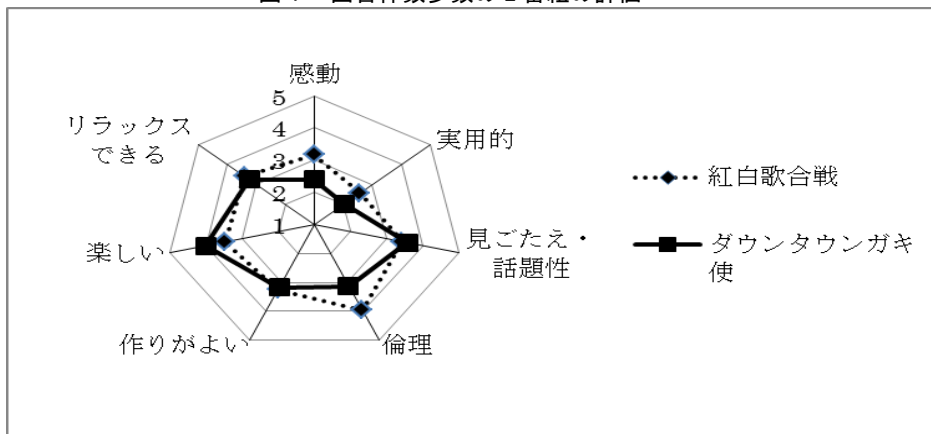
「楽しい」「リラックスできる」は独立しており、「感動できる」にはそれぞれ3項目、「倫理的に問題がある」には5項目、「実用的である」には3項目が、2つの調査について安定して同じ項目が同一クラスターに分類された。また、「作りが良い・見ごたえがある・話題性がある」の7項目については緩やかに結合して一つの大きなクラスターとなっており、さらに分割可能と見受けられる（両調査において含まれる項目に若干の違いがある）。そこで、「作りが良い・見ごたえがある・話題性がある」というクラスターについては、さらに「作りが良い」の4項目と「見ごたえ・話題性」の3項目の2つに分け、他の5つのクラスターと合わせ、最終的には7つのクラスターに分類した。

その後、評価結果によるクラスター分析を行った。つまり、それぞれの番組について、20項目の評価得点の平均値を算出し、続いて、7クラスターごとに振り分けられた項目の得点を合わせそれぞれの平均値を計算し、それをクラスターごとの番組得点として抽出した。番組得点からみる番組評価は、以下のように特徴づけられる。

(1) 多様なジャンルに適用するか（大晦日調査の結果から）

大晦日調査の結果をみると、27番組に対する回答が得られ、最も多くの回答が寄せられたのは『第64回紅白歌合戦』の452件で、次に『ダウンタウンのガキの使いやあらへんで!』が263件であった（番組評価結果については図1参照）。また、10件以上回答が集まった番組20番組について分析を行い、それらを分類したところ、「ニュース・情報」「バラエティ」「スポーツ」「音楽」の4つのジャンルに大別することができた。

図1 回答件数多数の2番組の評価



評価結果をジャンル別にみると、ジャンルごとの特徴が明らかになった（表1参照）。分析結果では、それぞれの番組に対し、見ごたえや話題性を感じながら感動を得て視聴する様子が窺える。さらに、ジャンル別の得点を比較してみると、「ニュース・情報」（3番組）は、倫理的問題はなく（得点：4.1）実用性も高く（得点：3.4）役立っているようだが、楽しく（得点：3.0）見ているわけではない。それに比べ「バラエティ」（7番組）は倫理得点が低めで（得点：3.7）実用性も低め（得点：2.8）だが、楽しく（得点：3.7）視聴している。また、「スポーツ」（2番組）も倫理得点はさほど高くはなく（得点：3.7）実用性は低く（得点：2.4）さらに、番組のつくり（得点：2.9）やリラックス性（得点：3.0）についても評価が低い。もっともリラックスして見ているジャンルは「音楽」（8番組）（得点：3.7）で、倫理的問題もなく（得点：4.3）とても楽しんでいる様子（得点：3.8）が窺える。番組としてのつくりのよさ（得点：3.6）も評価されている。

表1 7つのクラスターを基準とした4ジャンルごとの番組評価得点（平均値）と音楽ジャンルの詳細

ジャンル（番組数）	ニュース情報（3）	バラエティ（7）	スポーツ（2）	音楽（8）	音楽（クラシック）（3）	音楽（ポピュラー）（5）
楽しい	3.0	3.7	3.5	3.8	3.8	3.9
リラックスできる	3.4	3.4	3.0	3.7	3.9	3.6
感動できる	3.3	3.1	3.1	3.5	3.8	3.3
見ごたえがある・話題性がある	3.1	3.5	3.5	3.4	3.3	3.5
実用的である	3.4	2.8	2.4	2.8	3.0	2.7
作りが良い	3.2	3.3	2.9	3.6	3.6	3.5
倫理的に問題がある	4.1	3.7	3.7	4.3	4.5	4.1

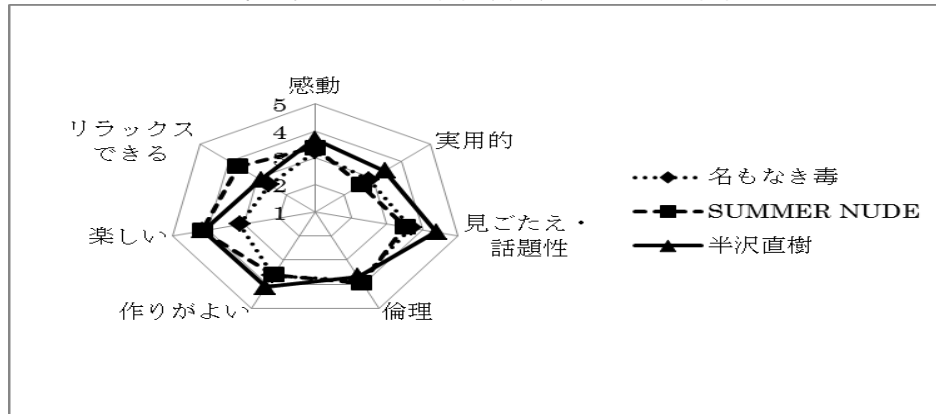
さらに、「音楽」のジャンルを「クラシック」（3番組）と「ポピュラー」（5番組）に分けた分析結果をみ

ると、表 1 に示す通り、クラシックは、より倫理的問題がなく（クラシック・ポピュラー順の得点：4.5、4.1）、リラックスしながらも（得点：3.9、3.6）、強い感動をもって（得点：3.8、3.3）享受している様子が窺える。また、教養が身につくなどの実用性が（得点：3.0、2.7）評価されていた。

(2) 同一ジャンルにおける番組評価は可能か（ドラマ調査の結果から）

ドラマという同一ジャンルについての調査結果をみると、大方のドラマは「作りが良」く「見ごたえ・話題性」があり「感動できる」番組となっており、「リラックス」して「楽しん」でいるが、「実用的」ではないという評価である。しかし、クラスター分析を行うと、同一ジャンルの番組でも、いくつかの評価結果から番組を分けることが可能だ。つまり、ドラマの設定状況や構成、ストーリーによっては、リラックスできず緊張して視聴する番組や、時には実用的と評価される番組があるなど、各々の番組により特徴がみられた（図 2 参照）。

図 2 番組群における回答件数多数の 3 ドラマの評価



たとえば、若者を対象とした『SUMMER NUDE』『ショムニ 2013』『スターマン・この星の恋』の番組評価をみると、ドラマは虚構の設定であるという認識からか、とりわけ実用性が低い（番組順の得点：2.6、2.7、2.5）がリラックスしながら（得点：3.7、3.6、3.8）とても楽しんでいる（得点：4.2、4.2、4.3）様子が窺える。一方、同様に実用性の低い番組（得点：2.9、2.7、2.5）として挙げられているのが『名もなき毒』『ガラスの家』『激流～私を覚えていますか？～』という番組群だが、事件・事故に絡みサスペンス要素が組み入れられている設定ゆえか、リラックスできず（得点：2.4、2.6、2.4）緊張した状況で視聴しているようだ。さらに、『半沢直樹』『七つの会議』というビジネス社会を舞台に繰り広げられる番組群は、「実用的である」（得点：3.4、3.1）とある一定の評価をしつつ緊張感を持ちながら（リラックス得点：2.9、2.7）楽しんでいる（得点：4.1、3.5）。構成や演出など作りの良さの得点（4.1、4.1）も高い。

5. 考察

評価結果によるクラスター分析によると、同一ジャンルの番組に対する評価は類似しており、ジャンル別に評価の特徴がみられていた。これは、つまり、多様な番組を評価するという点から考えると、ある程度の枠組みを提供することが可能となることを示している。さらに、ドラマというジャンルにおいてだが、同一ジャンルの番組でも、番組の特徴により評価が異なる程度分かった。これは、たとえ評価者数は少なくとも、人気のある番組と類似した項目についてほぼ同等の評価を得た番組は、今後、より多くの人に視聴してもらえる可能性を秘めているといえ、良質な番組を確保する契機となるのではないかと考えられる。

今後は、これらのクラスターにおける項目をより精査することにより、これまでの 20 項目から質問項目を大幅に圧縮することが可能となるだろう。ひいては、評価者への負担を軽減し、より回答しやすい調査へとつなげることができると考えられる。

さらに、今後の課題として、調査実施後、評価の速報値を提供できる体制づくり、すなわち、評価結果の集計を自動化することにより、筆者らによる独自調査だけでなく、共通の質問項目を利用した視聴者調査を汎用性の高いものとして確立していくことができよう。

参考文献

- 1) 小玉美意子 (2011): テレビ番組評価サイト Quae のマスコミュニケーション研究上の位置づけ, 『ソシオロジスト』No. 13, pp. 1-32.
- 2) 小玉美意子 (2013): Quae の活動 (武蔵大学 MMS 研究会), 日本マスコミュニケーション学会理論研究部会ワークショップ報告.
- 3) 戸田桂大・小玉美意子監修, 山下玲子編 (2013): 『ユーザーからのテレビ通信簿』, 学文社.
- 4) 武蔵メディアと社会研究会 (2005~2014): 『各年度活動報告書』, 武蔵大学総合研究所.