

Google Scholar/Books を用いた日本の 社会科学文献の引用分析

—— 格差社会論を事例として ——

山 本 耕 平

1 問題関心と研究の背景

社会学をはじめとするさまざまな学問分野における研究者の活動をそれぞれの分野の特徴にそくして把握することは、研究者の活動を記述・分析する科学社会学の一部としても、研究機関や研究者の研究評価を公正におこなう上でも、重要である（藤垣ほか 2004）。そこで注目される特徴の一つに、それぞれの分野においてどのようなタイプの文献がどの程度のインパクトを持つか、というものがある。論文や書籍などの文献は、研究者どうしが知見を伝えたり、他の研究者を説得したりするためのツールであり、各分野の活動の特徴を捉えるには、そうしたツールとしてどのような文献が重視されるかを分野ごとに把握することが有効である。一方で、社会学（もしくはより広く、社会科学）では相対的に書籍のほうが重視される傾向があることは、これまで繰り返し論じられてきた。その傾向は、成果を発表する手段として書籍が選ばれやすいという点でも（Bourke and Butler 1996; Kyvik 2003）、書籍が引用されやすいという点でも確認されている（Clements et al. 1995; Moed 2005; Small and Crane 1979）。上記の理念と、この現状認識との間には大きな溝がある。それは、社会科学において書籍が重視されるという現状があるにもかかわらず、書籍のインパクトを論文のそれと同じように評価する手法が未整備だということである（Kousha and Thelwall 2009; Tang 2008; Torres-Salinas et al. 2014）。この問題は、社会科学の内部でも論文が重視される世界と書籍が重視される世界が併存しているために（Cronin et al. 1997; Wolfe 1990）、社会科学者の活動を偏りなく把握することを妨げている。たとえば、アメリカの社会学部のランキングを分析した F. Moksony ら（2014）によれば、質的方法を専門とする教員が多いことはランクに負の効果をおよぼすが、その効果の一部は、質的方法を専門とする研究者は論文よりも書籍で成果を発表する傾向があり、ランキングが作成される際に書籍が業績として考慮されないことによって説明される。

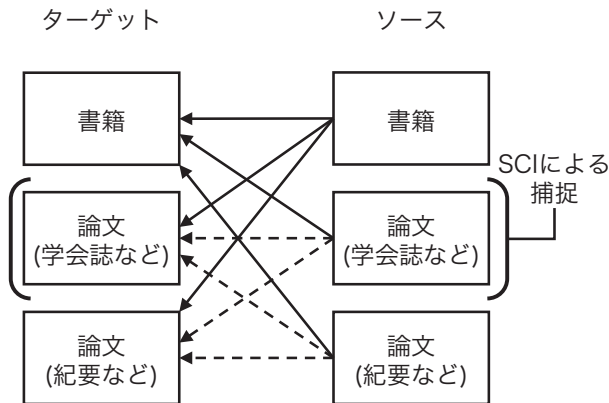


図1 引用のターゲットとソース

書籍のインパクトを測る試みはすでになされているが、文献のインパクトを測定するために従来もちいられてきた一般的な手法を日本の社会科学系書籍に適用することにはいくつかの問題がある。それらの問題を、引用される文献（ターゲット）の問題と、その文献を引用している文献（ソース）の問題に分けて整理する。

図1は、ターゲットとソース

をそれぞれ書籍、学会誌の論文、紀要などの論文に分け、引用・被引用の関係を図示している。実線の矢印は、書籍の影響力が相対的に強く、書籍をターゲットとした引用や書籍をソースとする引用が多いことを示している。「SCIによる捕捉」として括られている部分が、おおむね従来の手法によって捕捉されてきた範囲である。

まず、従来の引用分析は、ターゲットとして、論文、それも査読つき学会誌のような権威のあるジャーナルに発表される論文のみを対象とすることが多かった。これはおもに、引用分析でよく使われる Science Citation Index（以下、SCI）⁽¹⁾ が、そうしたジャーナルを中心にデータを収集しているためである。そのため、研究成果の発表が学会誌を中心に行われるような分野では、SCI において被引用数が多い文献がインパクトの大きい文献であると判断して差し支えないが、そうでない分野では、SCI に含まれない文献がよく引用されている可能性がある。実際、化学や生物学⁽²⁾ とは異なり、人文・社会科学では、他の方法で測った被引用数と SCI から得られるそれとがあまり一致しないことが指摘されてきた（Cronin et al. 1997; Moed 2005）。

次に、ソースも学会誌の論文に限定されることが多い。近年の事例でいえば、山本耕平と太郎丸博（2015）は、日英米の著名な学会誌の論文がどのような文献を参照しているかを検討しており、高橋大樹ほか（2017）も類似した観点から、『日本経営学会誌』に掲載

⁽¹⁾ 正確には、社会科学系の文献のデータは Social Science Citation Index に、人文系のデータは Arts & Humanities Citation Index に収録されているが、呼び分けが煩瑣なので以下は SCI で統一する。

⁽²⁾ 自然科学のなかでも分野間の差異は存在し、物理学ではプレプリント、コンピュータ・サイエンスではカンファレンス・ペーパーが引用されやすい、といった特徴がある（Kousha and Tehwall 2008）。

された論文をソースとして、日本の経営学の知的潮流を明らかにしようと試みている。また、兼子論と徳安彰（2012）は、日本での T. Parsons の受容にかんする通説を検証するにあたって、『社会学評論』に掲載された論文から Parsons の著作への引用を調査している。これらの場合には、ターゲットには書籍が含まれるが、ソースは従来どおり一部の学会誌に限定されている。しかし、図1に示すようにソースにもターゲットと同様の多様性があり、ソースを限定することによって、ある文献が持つ総合的なインパクトを測り損なっているかもしれない。日本の人文・社会科学では、学術誌以外で論文が掲載される紀要などの媒体のほか、『思想』（岩波書店）のような出版社が発行する雑誌の存在も見逃せないことを考えれば、学会誌を中心に文献のインパクトを測ることの問題はさらに大きいと言える⁽³⁾。

以上をまとめると、さまざまなソースからの引用を合わせると社会科学系の書籍が実際にどのくらい引用されているのか、そして、従来の研究では十分に注目されていない学会誌以外の論文や書籍からの引用はどの程度あり、どれくらい重要なのか、といった点はまだ明らかになっていない。そこで本稿では、Google が提供する各種サービスによって日本の社会科学系書籍の被引用数をソースの種類ごとに集計し、ソースを学会誌に限定しない場合に、書籍のインパクトがどのように測定されるのかを、格差社会論の書籍を事例として検討することにする。以下では、Google Scholar などの Google が提供するサービスを利用した引用分析の事例について説明し、日本の社会科学の文献にたいして類似した方法を適用することを検討する価値があることを確認する（2 節）。そして、本稿で採用する方法と分析対象について述べた後（3 節）、分析対象とする書籍の被引用数とその時系列的な変化、およびソースの種類ごとに見た被引用数と全体の被引用数との関連について報告する（4 節）。最後にまとめと今後の課題を述べる（5 節）。

2 先行研究

前節で述べた、SCI などの既存のデータベースをもちいて文献のインパクトを測ることの問題点は、そうしたインパクトの測定や、引用関係から研究者のネットワークを探索す

⁽³⁾ もう一点つけ加えると、既存のデータベースをもちいることには、言語の問題もある。自然科学に比べると、人文・社会科学では研究者が母国語で論文や書籍を書くことが多く、それらは SCI に収められにくい（Hicks 2004）。そのため、日本の社会科学において各種文献が持つインパクトを既存のデータベースによって測ろうとすると、学会誌以外の情報が欠如している上に日本語の文献にかんする情報も手薄である、という二重の問題が発生する。なお、近年は書籍のデータを収録した Book Citation Index（BKCI）も公表されているが、SCI 以上にデータが英語圏に偏っており、さらには Springer のような一部の出版社の書籍に偏っていてもいると報告されている（Torres-Salinas et al. 2014）。

るといった作業を精確におこなうことを第一義的な目標とする科学計量学においても重要な問題であり、代わりとなる手法の検討も主としてそれらの分野で検討されてきた。そこで有力な候補とされているのが、Google Scholar をはじめとする、Google が提供している検索サービスの活用である。

Google Scholar（以下、GS）とは、一言でいえば学術的な情報に特化した検索サービスであり、ウェブ上で公開されている論文や書籍、プレプリントなどのテキストやメタデータを検索することができる。どのように情報をアーカイブしているかは公表されておらず、少なくとも SCI のように対象となるジャーナルをあらかじめ選定するといった体系的な情報収集はなされていないようだが（Jacso 2005; Kousha and Thelwall 2008）、以下に見るように、既存のデータベースからは取得できない多様な引用にかんする情報が得られるため、引用分析における有効性に期待する研究者も多い。たとえば、社会学を含む 8 分野について、対象とした論文への引用件数を GS のデータによってカウントした K. Kousha and M. Thelwall (2007) によれば、人文・社会科学（社会学・経済学・教育学）の場合には SCI において漏れている引用が GS によって捕捉できるため、被引用数によって文献のインパクトを測るには GS のほうが有効である。また、J. Mingers and E. A. E. C. G. Lipitakis (2010) は、イギリスの経営学研究者の出版物への引用を Web of Science と GS で比較し、後者のほうがより多くの引用情報を取得できたと報告している。さらに、GS から得た引用のデータによって社会学系ジャーナルのランキングを作成している J. A. Jacobs (2016) によれば、SCI のデータを使った場合と比べてジャーナルの相対的なランキングは変化しないものの、被引用数が多い論文はトップランクのジャーナルに掲載されたものではなかった。これは、GS によって多様なソースからの引用を捕捉できることの効果であり、現在、日本の大学の多くがリポジトリで紀要論文を公開している現状を考えれば、日本でも同様に GS を使うことでより豊富なソースのデータが収集できると期待される。前述のとおり、Google がどのように情報を収集しているのか公開されていないことから、信頼性の面で GS は劣るとされることもあるが、A. A. M. Prins et al. (2016) によれば、情報源がはっきりしない検索結果（彼らの報告によれば、検索結果の 10% ほど）を除外するなどのデータクリーニングをおこなえば、研究評価にも十分に使えるという。

Google Books（以下、GB）は、Google がスキャンした書籍の全文もしくは一部を公開し、内容を検索できるようにしているサービスである。これを使用することで、ある文献を引用している書籍を調べることができる。Kousha and Thelwall (2009) は GB をもちいて、論文への書籍からの引用をカウントし、SCI から取得できる論文からの引用を補完することを提案している。GB はまた、GS を補完する手段として使うこともできる。GS にも書

表 1 J-STAGE と GS の比較 (2017 年 8 月 31 日時点)

J-STAGE	Google Scholar
	Baldwin and Allison 2015
福井 2015	福井 2015
	林 2012
	Kamo 2012
Kikkawa and Fujihara 2012	Kikkawa and Fujihara 2012
	Kojima 2013
Kojima 2015	Kojima 2015
仲・前田 2014	仲・前田 2014
Takahashi 2016	Takahashi 2016
	山本ほか 2015

籍の情報は含まれるが、GS が扱う文献データは学術的なものに限られるため、日本では例えば各種新書や総合雑誌のように、場合によっては非学術的とみなされる文献が GS から漏れてしまう可能性があるからである。

GS と GB を併用した研究の例として、A. Abrizah and Thelwall (2014) は、非英語圏で出版された書籍のインパクトを測ることを試みている。具体的には、マレーシアの大学系出版会から刊行されている書籍をサンプリングし、それらのタイトルや著著名などを GS と GB で検索した結果から、よく引用される文献のランキングを作成しており、そのランキングに見られる傾向が先行研究の知見とも整合的であったとして、GS や GB を使用することには一定の有効性があると述べている。このようにして彼らは、非英語圏の文献にかんする情報を得る手段として GS と GB を位置づけているが、同時に、日本語や中国語といったテキスト化のエラーが起こりやすい言語については注意が必要である、とも述べている。この点については次節で検討する。

先行研究の検討からは離れるが、日本語の文献についても、上記と同様に GS のほうが引用にかんする情報を豊富に有することが言えそうであることを、具体的な事例をもちいて確認しておきたい。日本の場合、近年では、科学技術振興機構が運営している J-STAGE において、公開されている論文の被引用文献が一覧表示されるようになっていいる。しかしこれは、同じように J-STAGE で公開されている文献からの引用しか登録されないため、論文以外のタイプの文献からの引用は原理的に登録されないし、紀要などで発表された論文からの引用も対象外である。つまり、図 1 で示した SCI と同様の問題があると予想される。そこで一例として、J-STAGE で公開されている有田伸「比較を通じてみる東アジアの社会階層構造」(2009 年、『社会学評論』59 (4)) を引用している文献を、

J-STAGE の被引用文献リストと GS によってそれぞれ検索した結果をまとめたものが表 1 である。ここから分かるように、J-STAGE では表示されないが GS では表示されるものが 5 件ある。そのうち、Baldwin and Allison は書籍であり、山本ほかは本誌『京都社会学年報』で発表された論文である。このように、被引用数によって文献のインパクトを知るための手段としては、J-STAGE のデータもいまだ十分な情報量を有しているとは言えず、それと比較すると GS による検索は、もっとも多様な文献からの引用を集めることができるという点で、現時点では最良の手段だと思われるのである。

3 方法とデータ

3-1 方法

本稿では、前節で述べたように GS を主な手段とし、補完的に GB ももちいて、日本の社会科学における書籍のインパクトがどのように測定できるかを検討する。具体的には、次節で述べる 7 冊の書籍をターゲットとして、それらを引用しているソースの数を種類別に数える。図 2 に示すように、GS で表示される検索結果の単位は 1 件の文献であり、その文献が何らかの文献から引用されている場合、「引用元」(図 3 では Cited by) という項目が左下に表示される。ここから、その文献を引用している文献の一覧を見ることができる(図 3)。この一覧には、テキスト化されていない文献でも(おそらくは文字の形状を認識して)含まれており、検索の精度は高いと考えられる。図 2 に示しているのは論文の場合だが、書籍についても同様に、当該書籍のデータが Google によってアーカイブされている場合、同じように「引用元」が表示される。そうでない場合は、書名を GS で検索することで、その書名を本文中に含む文献を検索した。この場合、当該の書名を含む文献が

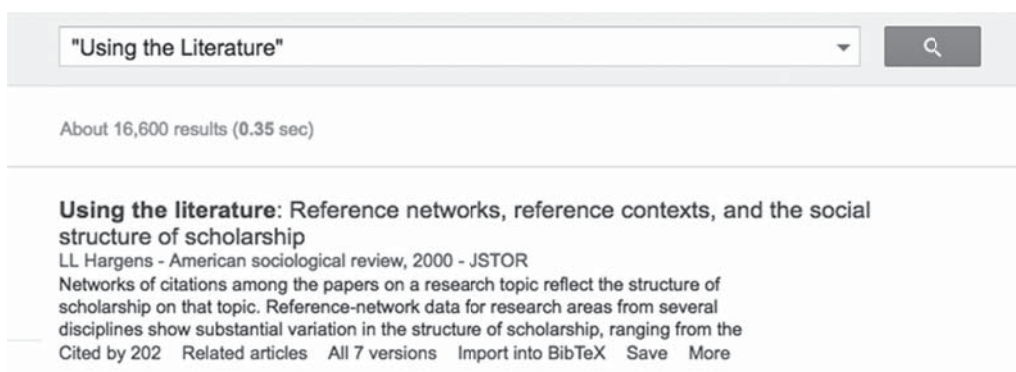


図 2 Google Scholar の検索結果の例 (1)

Using the literature: Reference networks, reference contexts, and the social structure of...

☐ Search within citing articles

[CITATION] Research genres: Explorations and applications

[PDF] researchgate.net

J Swales - 2004 - Ernst Klett Sprachen

Cited by 2287 Related articles All 7 versions Import into BibTeX Save More

[BOOK] Social network analysis

D Knoke, S Yang - 2008 - books.google.com

In the thoroughly updated edition of Social Network Analysis, authors David Knoke and Song Yang take into account the vast number of changes in the field that have occurred in the 25 years since the first edition was published. They cover various issues in basic

Cited by 1352 Related articles All 3 versions Import into BibTeX Save More

The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999

[PDF] researchgate.net

J Moody - American sociological review, 2004 - journals.sagepub.com

Has sociology become more socially integrated over the last 30 years? Recent work in the sociology of knowledge demonstrates a direct linkage between social interaction patterns and the structure of ideas, suggesting that scientific collaboration networks affect scientific

Cited by 817 Related articles All 15 versions Import into BibTeX Save More

What do citation counts measure? A review of studies on citing behavior

[PDF] mpg.de

L Bornmann, HD Daniel - Journal of documentation, 2008 - emeraldinsight.com

Purpose-The purpose of this paper is to present a narrative review of studies on the citing behavior of scientists, covering mainly research published in the last 15 years. Based on the

図3 Google Scholar の検索結果の例 (2)

すべてヒットするため、検索結果や全文に目を通し、false match がないかどうか確認した。また、「引用元」が表示される場合でも、その文献を引用している文献が網羅されているわけではない。たとえば、『日本労働研究雑誌』の論文はすべてテキスト化されて公開されているため、大竹文雄『日本の不平等』を引用している論文で『日本労働研究雑誌』に収録されているものは、原則としてはすべて「引用元」の一覧に表示されるはずである。しかし、『日本労働研究雑誌』に発表されている佐藤嘉倫・吉田崇「貧困の世代間連鎖の実証研究」は、同書を引用しているにもかかわらず、『日本の不平等』の「引用元」一覧には表示されない。そこで、「引用元」から表示される文献のリストを作成した後に、書名で検索をおこない、重複しないものをリストに加えていった。

GB については、対象となる書籍のタイトルや著者名を検索することが基本となるが、Abrizah and Thelwall (2014) が注意を促すように、日本語の文献はテキストデータへの変換の際にエラーが起りやすいかもしれない。GB では、ページのスキャンデータは公開されていなくとも、テキストデータをスニペット表示⁽⁴⁾できる文献も多いが、こちらの場合はさらに誤字や文字化けが多い。そのため、検索ワードを工夫することで、できるだけ漏れをなくすよう工夫した。たとえば、荻谷剛彦『階層化日本と教育危機』を引用している文献を検索する場合、「荻谷 教育危機」といったように、複雑な漢字をできるだけ避けると同時に書名のキーワードを残した検索ワードを用いることで、本来なら『階層化

(4) 検索結果が一覧で表示される際に、検索語を含むページの一部が表示されること。

『日本と教育危機』を引用している文献が文字化けのせいでヒットしない、という事態を防ぐことを試みた。

以上のように、ターゲットの書籍を引用している文献の情報を集めた後、それらの文献の数を単純にカウントすることに加えて、発行年と文献のタイプによって分類した。この作業の目的は、すでに述べたとおり、学会誌を中心として書籍の被引用数を数えた場合と、他のソースも加えた場合とで、どのくらい傾向が異なるのかを確認することである。したがって、文献のタイプは、書籍と商業誌、学会誌に発表された論文とそれに準ずるもの、学会誌以外の論文（紀要論文やワーキング・ペーパー）、その他（博士論文や学会報告の要旨など）、の四つに分類した。学会誌以外の論文のなかには、本誌のように査読が行われているものもあればそうでない紀要論文も含まれるし、ワーキング・ペーパーの類もこの分類に含まれる。これらの論文は、SCIのような既存のデータベース（あるいは、今後開発されるかもしれない類似のデータベース）には含まれない（であろう）点で共通したもののとして、本稿ではひとまとめに扱っている。

3-2 データ

ターゲットとして分析対象とするのは、2000年代に日本でブームになったとされる「格差社会論」に関連する書籍である。本稿で調査したいのは人文・社会科学系の書籍のインパクトだが、GSをもちいた引用のカウントは、あまり古い文献にたいしては有効ではないおそれがある（Bauer and Bakkalbasi 2005）。その点、格差社会論が流行したとされるのはおおむね2000年以降なので、関連した書籍を引用している文献も当然2000年代以降のものであり、電子化されて公開されている確率が高く、GSやGBをもちいることである程度の量の引用が捕捉できると期待される。また、格差社会論のブームはすでに落ち着きを見せており、被引用数の時系列的な変化を見ることにも適している。

具体的には、まずCiNii Booksで、1995年から2012年まで⁽⁵⁾に発行された日本語の書籍を対象とし、「格差社会」と「下流」をそれぞれキーワード検索して得られた結果から、所蔵館が多いもの（2017年8月31日時点）から順に、「格差社会」で5点、「下流」で2点を選択した。これに、格差社会論を巻き起こした書籍としてしばしば言及される（cf. 平沢 2014; 盛山 2011）佐藤俊樹『不平等社会日本』（2000年、中公新書）を加えた以下の8点から、被引用数が極端に少ない藤田編（被引用数25、表2参照）をのぞく7点を分析対

⁽⁵⁾ 2012年までに限定したのは、それ以降に発行された書籍はまだ被引用数のピークを迎えていない可能性があり、時系列的な変化を見るのに適していないと考えたからである。

象とした。

- 橘木俊詔, 2006, 『格差社会——何が問題なのか』 岩波新書. (所蔵館 626 館)
- 荏谷剛彦, 2011, 『階層化日本と教育危機——不平等再生産から意欲格差社会へ』 有信堂高文社. (所蔵館 543 館)
- 藤田英典編, 2007, 『誰のための「教育再生」か』 岩波新書. (所蔵館 535 館)
- 山田昌弘, 2004, 『希望格差社会——「負け組」の絶望感が日本を引き裂く』 筑摩書房. (所蔵館 531 館)
- 大竹文雄, 2005, 『日本の不平等——格差社会の幻想と未来』 日本経済新聞社. (所蔵館 485 館)
- 三浦展, 2005, 『下流社会——新たな階層集団の出現』 光文社新書. (所蔵館 556 館)
- 内田樹, 2007, 『下流志向——学ばない子どもたち働かない若者たち』 講談社. (所蔵館 498 館)
- 佐藤俊樹, 2000, 『不平等社会日本——さよなら総中流』 中公新書. (所蔵館 474 館)

一般的な引用分析では 1,000 本を超える論文を分析の対象にすることも多く、7 冊という数はかなり少ないが、前節で触れたように GS の検索結果にはある程度のクリーニングが必要であり、検索にヒットした文献を一つずつチェックするので、あまり数を増やすことは得策ではないと思われる。書籍への引用を集計・分析する研究では、1 冊から十数冊の本への引用をカウントしている事例もある (cf. Bar-Ilan 2010; Porta et al. 2006)。なお、「格差社会」の書籍については、大竹『日本の不平等』に次いで所蔵館数が多いのは佐藤嘉倫・尾嶋史章編『格差と多様性』(2011 年、東京大学出版会) だが、その所蔵館数は 378 であり、『日本の不平等』とは大きな差がある。「下流」についても、内田『下流志向』に続くのは竹内洋『大学の下流化』(2011 年、NTT 出版) の 322 館で、やはり大きな差がある。上に挙げた 8 点を、広く読まれたという意味で「代表的」な格差社会論とすることは、さほど問題ではないと思われる。また、この 8 冊のなかには、専門的な内容を含む荏谷と大竹、一般読者向けの橘木や山田、一般読者向けではあるが研究者の間でも論争を惹起した佐藤、学術的な論調とは言えない三浦と内田など、内容から見てもいくつかのタイプの格差社会論が含まれており、その点でも一定の代表性が確保できるものと判断した。

4 結果

4-1 ソースの内訳と時系列的变化

対象とした8冊（ここでは参考のため、藤田編『誰のための「教育再生」か』も含めてある）の書籍の被引用数の合計と内訳を表2に示した。「格差社会」をキーワードとした4冊と佐藤は、紀要などからの引用が半分程度かそれ以上であり、学会誌からの引用は最大でも25%をやや上回る程度である。言い換えれば、学会誌の論文のみをソースとしてこれらの書籍のインパクトを測った場合、高く見積もっても実態の1/3ほどしか捉えられないかもしれない、ということである。さらに学会誌からの引用と紀要からの引用との差が大きいのが「下流」をキーワードとする2冊で、これらの書籍は、学会誌ではほとんど引用されていないが、それに比べると紀要ではよく引用されている。とくに三浦は、紀要からの被引用数は橋木を上回っており、「下流社会」という言葉の流行をアカデミズムの外の出来事と捉えることはできないことを示唆している。

ターゲットごとに、ソースの内訳とともに被引用数の時系列的な変化を表したものが図4～10である。多くの書籍は、出版から1～2年の内に被引用数のピークを迎えている点で共通している。BKCIのデータをもちいてさまざまな分野における書籍の引用のされ方を比較したR. Tang (2008)によれば、宗教学、歴史学、経済学といった人文・社会科学系の分野では、書籍の被引用数は平均的には出版から2年目にピークを迎えており、他の分野よりも短い。本稿の結果はこれと整合的であり、GSやGBから得られる被引用数のデータがさほど偏っていないことを示唆する結果といっていよい。さらに、荻谷や佐藤のように

表2 被引用数の内訳

橋木 2006				荻谷 2001			
計		125 (100%)		計		297 (100%)	
書籍	40 (32%)	学会	21 (16.8%)	書籍	13 (4.4%)	学会	79 (26.6%)
紀要	62 (49.6%)	他	2 (1.6%)	紀要	186 (62.6%)	他	19 (6.4%)
藤田編 2007				山田 2004			
計		25 (100%)		計		254 (100%)	
書籍	6 (24%)	学会	1 (4%)	書籍	16 (6.3%)	学会	72 (28.3%)
紀要	18 (72%)	他	0 (0%)	紀要	156 (61.4%)	他	10 (3.9%)
大竹 2005				三浦 2005			
計		205 (100%)		計		169 (100%)	
書籍	40 (19.5%)	学会	52 (25.4%)	書籍	67 (39.6%)	学会	17 (10.1%)
紀要	110 (53.7%)	他	3 (1.5%)	紀要	83 (49.1%)	他	2 (1.2%)
内田 2007				佐藤 2000			
計		87 (100%)		計		264 (100%)	
書籍	15 (17.2%)	学会	4 (4.6%)	書籍	79 (29.9%)	学会	63 (23.9%)
紀要	64 (73.6%)	他	4 (4.6%)	紀要	112 (42.4%)	他	10 (3.8%)

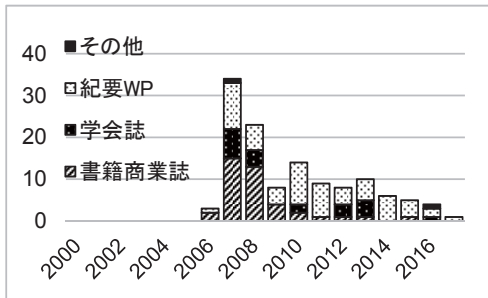


図 4 橋木 (2006) の被引用数

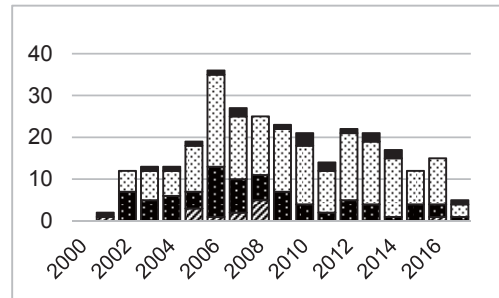


図 5 荻谷 (2001) の被引用数

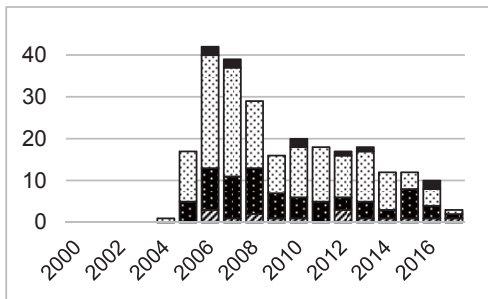


図 6 山田 (2004) の被引用数

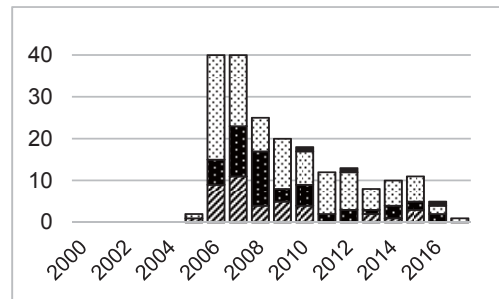


図 7 大竹 (2005) の被引用数

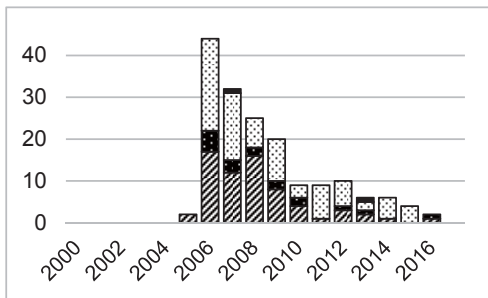


図 8 三浦 (2005) の被引用数

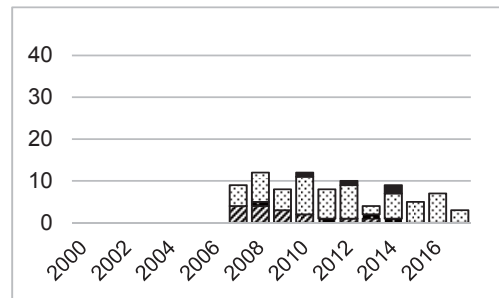


図 9 内田 (2007) の被引用数

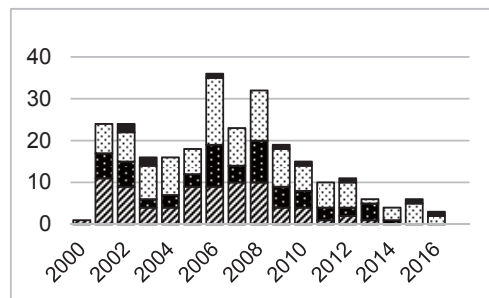


図 10 佐藤 (2000) の被引用数

2000 年代初頭に出版された書籍の時系列的な変化を見ると、2006 ～ 7 年にかけて被引用件数が上昇しており、アカデミズムのなかでも格差社会論のブームがあったことがうかがえる⁽⁶⁾。荻谷、山田、大竹のように、被引用件数のピークを過ぎてからも、おもに紀要からある程度の数で引用されている書籍もあるが、こうした通時的な動向も、学会誌からの引用を見るだけでは把握できないものである。

4-2 被引用数にたいする寄与の分析

さらに、それぞれのターゲットの被引用数にたいしてどのタイプのソースからの引用が相対的に強い影響を与えているかを確認するために、さらなる 2 つの分析をおこなう。まず、ソース別に被引用数の平均と標準偏差（表 3）、および、ソースごとの被引用数どうし、および総引用数との相関係数を求めた（表 4）。学会誌・紀要等・その他はお互いに、また総引用数ともよく相関しており統計的にも有意だが、書籍からの被引用数はどの変数とも相関していない。散布図（非表示）を見ると、総引用数が多い荻谷や山田にたいする書籍からの引用が少ないため、相関が弱くなっている。7 ケースしかないため一般化はできないが⁽⁷⁾、この分析結果を正しいとするならば、書籍のインパクトを測定するにあたって他の書籍からの引用を考慮する必要性は低い、と考えられる。

表 3 に示した結果によれば、分散がもっとも大きいのは紀要等なので、紀要等からの引用が総引用数にたいして相対的に大きな影響を与えていると予想できる。そこで第二の分析として、各ターゲットの被引用数の時系列的な変化にたいして、各ソースからの引用がどのていど寄与しているのかを寄与度によって表してみよう。あるターゲットの出版後 j 年から $j+1$ 年（ $j = 1, 2, 3, 4$ ）の被引用数の変化率に占めるソースの種類 i （ $i = 1, 2, 3$ 。その他は除く）の寄与度 R_{ij} は、GDP 成長率を民間需要や公的需要の寄与度に分解する際に用いられるものと同様の下式によって得られる。ただし、 C_{ij} は時点 j におけるソース i からの被引用数である。

$$R_{ij} = \frac{(C_{ij+1} - C_{ij})}{\sum_{i=1}^3 C_{ij}}$$

⁽⁶⁾ 佐藤の被引用数が 2008 年にふたたび上昇しているが、これは 2005 年 SSM のデータをもちいた文献が発表され、1995 年 SSM のデータにもとづいて議論を展開していた佐藤に言及する文献が増えた時期と一致している。

⁽⁷⁾ また、次節で述べるように、日本語の書籍はあまり GB に登録されていないため、セレクション・バイアスが生じている可能性にも留意が必要である。

表 3 被引用数の平均と標準偏差

	書籍	学会誌	紀要等	その他
平均	38.6	44.0	110.4	7.1
標準偏差	26.3	29.7	46.6	6.3

表 4 総引用数およびソース別被引用数の相関係数 (N=7)

	総数	書籍	学会誌	紀要等
書籍	0.05			
学会誌	0.96***	-0.13		
紀要等	0.91**	-0.36	0.92**	
その他	0.81*	-0.34	0.80*	0.89**

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

寄与度の正負の向きが変化率と同じ方向である場合、被引用数の変化をその方向に押し進める効果をもっており、反対の方向である場合は、被引用数の変化を逆方向に進める効果を持っている。たとえば、変化率が正であるときに学会誌の寄与度が正、書籍の寄与度が負であれば、学会誌からの引用が増えたことによって被引用数が増加したが、一方で書籍からの引用が減ったことによってその増加が抑制された、というように判断できる。ターゲットごとに、被引用数の年ごとの変化率と上式にしたがって算出された寄与度を図示したものが、図 11～17 である。すべてのターゲットにおおむね共通して見出されるのは、学会誌の論文の寄与度が低いのにたいして、紀要等の寄与度は全般的に高くなっていることである。たとえば、橘木や内田の被引用数は、書籍と紀要からの引用が減少することによって全体としても減少し、それらからの引用が増えることによって押し上げられている。荻谷と大竹は一部の時点において学会誌の寄与度が大きくなっており、専門的な内容を含む書籍として他の 5 冊とは異なる特徴を示してはいるが、それでもなお、紀要等の寄与度は無視できない大きさである。これらの結果は、本稿で検討した書籍のインパクトの大小が、学会誌の論文以外の媒体からの引用によって左右されていることを示唆している。出版年から 2017 年までの被引用数の総計で見た表 4 の結果からは、学会誌からの引用も紀要等からの引用も総引用数と強く関連しており、被引用数で測ったインパクトの指標としては学会誌からの引用を見るだけでも十分かもしれないという可能性も考えられるが、図 11～17 の結果のように時系列的な変化を考慮に入れると、紀要等からの引用もやはり考慮する必要があると言えるだろう。

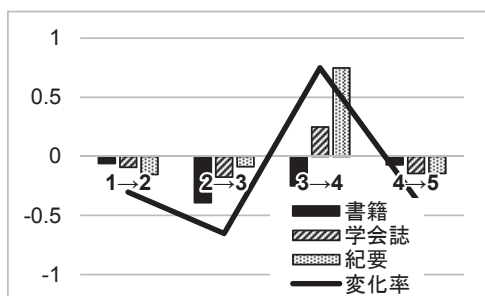


図 11 橋本 (2006) の変化率

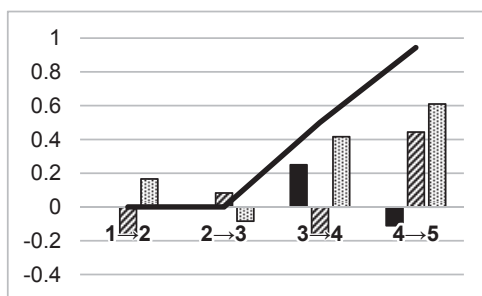


図 12 荻谷 (2001) の変化率

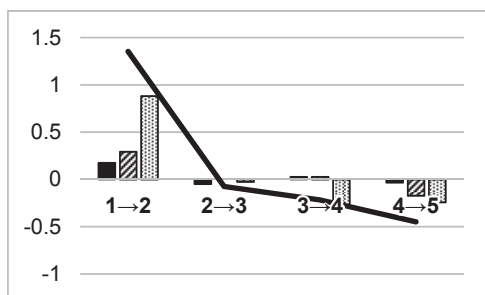


図 13 山田 (2004) の変化率

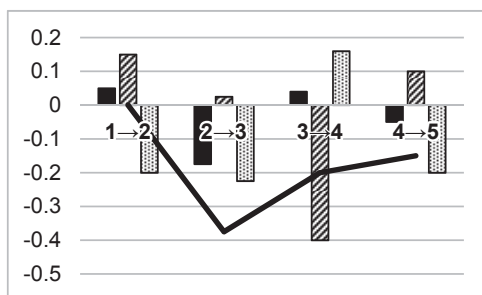


図 14 大竹 (2005) の変化率

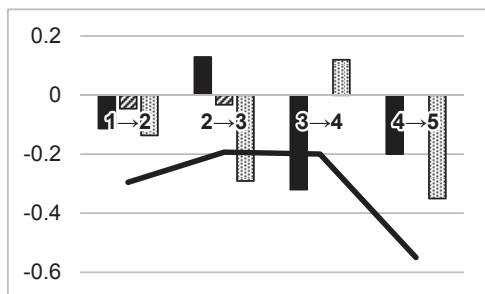


図 15 三浦 (2005) の変化率

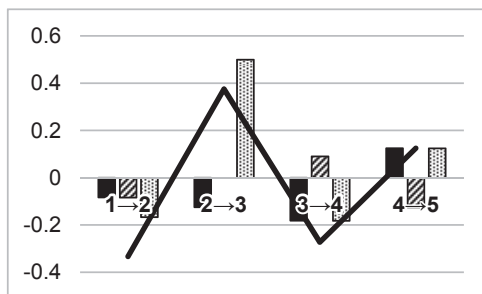


図 16 内田 (2007) の変化率

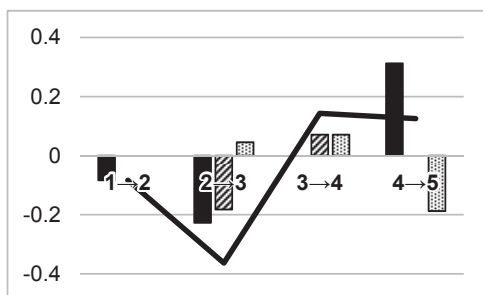


図 17 佐藤 (2000) の変化率

5 議論

前節で確認した結果から、日本の社会科学系書籍のインパクトを測るにあたって、少なくとも学会誌とともに紀要等からの引用は考慮する必要があることが分かった。引用分析のツールとしての有効性についていえば、少なくとも GS を利用することには有効性があると言える。おもに GB から情報を得られる書籍からの引用は、総引用数と関連しておらず、ターゲットである書籍のインパクトの指標に含める必要性は低いかもしれない。一方で、GS から情報を得られる紀要等からの引用は、総引用数と強く関連しているのみならず、総引用数の時系列的な変化への寄与度も高く、書籍のインパクトにたいして無視できない影響力を持っていると考えてよい。GS をもちいて、既存のデータベースでは捕捉できなかった多様なソースからの引用をカウントし、さらにソースの種類を分類した本稿のアプローチによって、これまで十分に明らかにされてこなかった書籍のインパクトの規定要因について、その一端を明らかにすることができた。さらなる発展の可能性として、2 節で触れたように、本稿で書籍への引用について確認したことと同様のことは、程度の差こそあれ論文への引用についても言えるので、論文のインパクトを測る際にも、GS を利用することで、どのような文献がどの程度・どのように引用されているのかをより正確に把握できるだろう。それによって、よく引用される書籍とよく引用される論文で引用のされ方にパターンの違いがあるかどうかなどを明らかにする道が開かれれば、書籍による成果の出版と論文によるそれという二つの「出版文化」(Moksony et al. 2014) が存在するかどうか、それらが(存在するとして)どの程度たがいに乖離したものなのか、といった科学社会的に興味深いリサーチ・クエスチョンに導かれることになるだろう。GS の検索結果を一括ダウンロードできるソフトウェアは開発されており本稿の調査でももちいたものの(Harzing 2010)、3 節で述べたようにクリーニングが必要なので、効率の点で問題は残る⁽⁸⁾が、今後こうした引用分析の試みが増えれば効率性の問題は解決されていくかもしれない。

ただし、そうした効率性の問題が解決されても、課題も多い。本稿では、本文の中に当該書籍への言及があればすべて等しく 1 件の引用とみなしているが、その書籍のインパクトを測りたければ、どのような文脈で引用されているかを確認する必要があるだろ

⁽⁸⁾ 3 節で述べたように、当該文献の「引用元」にはその文献を引用しているソースが網羅されているわけではないので、「引用元」に含まれないソースを逐一検索する作業も必要である。それがどの程度であるかは文献によって異なるが、本稿で調査した文献の場合、多い場合で 50 件程度が「引用元」から漏れていた。

う (McCain and Turner 1989)。これは、どのようなデータベースをもちいるとしても引用分析にはつねについてまわる問題ではあるが、テキストの内部を容易に検索できる GS の利点を活かして解決を図るべき問題かもしれない。また、繰り返し述べていることではあるが、データベースの信頼性の問題も解決されているわけではない。今回の結果ではどの書籍にたいしても書籍からの引用が少ないが、英語圏の先行研究では GB をもちいて書籍から論文への被引用件数を数えても相当な数の引用が見つまっている (Kousha and Thelwall 2009) ことからすると、GB に十分な量の日本語書籍が登録されていないことが要因だと考えられる。もし GB に登録される日本語書籍に偏りがあるならば、本稿でもちいたデータにもセレクション・バイアスがかかっている可能性がある。現時点では、書籍からの引用についてはカウントせず、学会誌と紀要等の論文からの引用に絞って書籍のインパクトを測定することが最適かもしれない。

このように、引用分析そのものに関わる課題も、GS や GB に特有の課題も残っており、GS や GB が既存のデータベースに取って代わる引用分析のツールになりえるかどうかは研究者の間でも係争中であるが、本稿の結果は、文献のインパクトを測るにあたって学会誌の論文に偏重していた従来の手法を乗り越えるための一案として、GS や GB を活用することを前向きに検討すべきであることを示している。

文献

- Abrizah, A. and M. Thelwall, 2014, "Can the Impact of Non-Western Academic Books be Measured? An Investigation of Google Books and Google Scholar for Malaysia," *Journal of the Association for Information Science and Technology*, 65 (12): 2498-508.
- Bar-Ilan, J., 2010, "Citations to the 'Introduction to Informetrics' Indexed by WOS, Scopus and Google Scholar," *Scientometrics*, 82: 495-506.
- Bauer, K. and N. Bakkalbasi, 2005, "An Examination of Citation Counts in a New Scholarly Communication Environment," *D-Lib Magazine*, 11 (9). (2017 年 9 月 1 日取得, <http://www.dlib.org/dlib/september05/bauer/09bauer.html>)
- Bourke, P. and L. Butler, 1996, "Publication Types, Citation Rates and Evaluation," *Scientometrics*, 37(3): 473-94.
- Clements, E. S., W. W. Powell, K. McIlwaine and D. Okamoto, 1995, "Careers in Print: Books, Journals, and Scholarly Reputations," *American Journal of Sociology*, 101 (2): 433-94.
- Cronin, B., H. Snyder and H. Atkins, 1997, "Comparative Citation Rankings of Authors in Monographic and Journal Literature: A Study of Sociology," *Journal of Documentation*, 53 (3): 263-73.
- 藤垣裕子・平川秀幸・富澤宏之・調麻佐志・林隆之・牧野淳一郎, 2004, 『研究評価・科学論のための科学計量学入門』丸善.
- Harzing, A. W., 2010, *Publish or Perish Book: Your Guide to Effective and Responsible Citation Analysis*, Melbourne: Tarma Software Research.
- Hicks, D. M., 2004, "The Four Literatures of Social Science," H. Moed, W. Glänzel and U. Schmoch eds., *Handbook of Quantitative Science and Technology Research*, Dordrecht: Kluwer Academic, 473-96.
- 平沢和司, 2014, 『格差の社会学入門——学歴と階層から考える』北海道大学出版会.

- Jacobs, J. A., 2016, "Journal Rankings in Sociology: Using the H Index with Google Scholar," *American Sociologist*, 47: 192-224.
- Jacso, P., 2005, "As We May Search: Comparison of Major Features of the Web of Science, Scopus, and Google Scholar Citation-based and Citation-enhanced Databases," *Current Science*, 89 (9): 1537-47.
- 兼子論・徳安彰, 2012, 「タルコット・パーソンズの日本社会学における受容過程の分析——従来のパーソンズ受容の認識に対する『社会学評論』調査による応答」『社会志林』58 (4): 143-162.
- Kousha, K. and M. Thelwall, 2007, "Google Scholar citations and Google Web/URL citations: A Multi-Discipline Exploratory Analysis," *Journal of the American Society for Information Science and Technology*, 58 (7): 1055-65.
- , 2008, "Sources of Google Scholar Citations outside the Science Citation Index: A Comparison between Four Science Disciplines," *Scientometrics*, 74 (2): 273-94.
- , 2009, "Google Books Search: Citation Analysis for Social Science and the Humanities," *Journal of the American Society for Information Science and Technology*, 60 (8): 1537-49.
- Kyvik, S., 2003, "Changing Trends in Publishing Behaviour among University Faculty, 1980-2000," *Scientometrics*, 58 (1): 35-48.
- McCain, K. W. and K. Turner, 1989, "Citation Context Analysis and Aging Patterns of Journal Articles in Molecular Genetics," *Scientometrics*, 17 (1-2): 127-63.
- Mingers, J. and E. A. E. C. G. Lipitakis, 2010, "Counting the Citations: A Comparison of Web of Science and Google Scholar in the Field of Business and Management," *Scientometrics*, 85: 613-25.
- Moed, H. F., 2005, *Citation Analysis in Research Evaluation*, Dordrecht: Springer.
- Moksony, F., R. Hegedus and M. Csaszar, 2014, "Rankings, Research Styles, and Publication Cultures: A Study of American Sociology Departments," *Scientometrics*, 101: 1715-29.
- Porta, M., E. Fernandez and E. Puigdomènech, 2006, "Book Citations: Influence of Epidemiologic Thought in the Academic Community," *Rev. Saúde Pública*, 40: 50-6.
- Prins, A. A. M., R. Costas, T. N. van Leeuwen and P. F. Wouters, 2016, "Using Google Scholar in Research Evaluation of Humanities and Social Science Programs: A Comparison with Web of Science Data," *Research Evaluation*, 25 (3): 264-70.
- 盛山和夫, 2011, 「階層的不平等研究の最近の動向と課題」『海外社会保障研究』(177): 52-64.
- Small, H. G. and D. Crane, 1979, "Specialties and Disciplines in Science and Social Science: An Examination of Their Structure Using Citation Indexes," *Scientometrics*, 1 (5-6): 445-61.
- 高橋大樹・積田淳史・渡部博志, 2017, 「経営学のヒストリカル・レビューに向けて——引用分析およびテキスト分析」『武蔵野大学政治経済研究所年報』(14): 97-120.
- Tang, R., 2008, "Citation Characteristics and Intellectual Acceptance of Scholarly Monographs," *College & Research Libraries*, 69 (4): 356-69.
- Torres-Salinas, D., N. Robinson-Garcia, A. Cabezas-Clavijo and E. Jimenez-Contreras, 2014, "Analyzing the Citation Characteristics of Books: Edited Books, Book Series and Publisher Types in the Book Citation Index," *Scientometrics*, 98: 2113-27.
- Wolfe, A., 1990, "Books vs. Articles: Two Ways of Publishing Sociology," *Sociological Forum*, 5 (3): 477-89.
- 山本耕平・太郎丸博, 2015, 「社会学の方法と引用文化の日英米比較」『理論と方法』30 (2): 165-80.

(やまもと こうへい・特任研究員)