

はじめに

プログラム評価とは、端的に言うと字義通り、「プログラム」を「評価」することである。プログラム評価の発展背景、意義や有用性、方法論、評価者の役割などは、医学教育領域では馴染みがない部分も多い¹⁾。そこで本セクションでは、①プログラム評価の発展背景と定義、②代表的な評価形式であるプロセス評価とアウトカム評価、③評価者の役割について概説した上で、今後のプログラム評価研究においてその必要性が高いと考えられる④アカウンタビリティ向上と⑤評価キャパシティ構築の視点について検討する。

1. プログラム評価とは何か：発展背景と定義

プログラム評価は、心理学、教育学、社会学といった社会科学領域の研究者が中心となって、主に米国において発展してきた。発展初期である1950年代～1970年代は、実験心理学の手法^{2,3)}を用いた社会プログラム・公共サービスの評価の方法論が整備された時期であった。特に、実験・準実験デザインを用いたデータに基づく公共プログラムや社会サービスの変革(reform)の必要性が指摘された⁴⁾。しかし、そのような社会実験(experimenting society)を応用したプログラム評価は、政治的・社会経済的要因の影響が常に存在するフィールドでは、非現実的かつ視野狭窄であるという見方が専門家の間で主流となっていった⁵⁾。その後、学際的な領域の研究者により、実用重視評価⁶⁾、エンパワメント評価⁷⁾、理論主導型評価⁸⁾といった様々なアプローチが開発され、発展してきた。

プログラムとは「何らかの問題解決や目標達成を目的に人が中心となって行う実践的介入」のことを指し、一般に、①地域や組織ですでに起きている問題・課題の解決(問題解決)、②そもそもの問題自体の発生の予防や生活の質やウェルビーイングの促進、③様々な学びやスキルの習得を目的とした教育やトレーニング(訓練)の3つの目的に大別できる⁹⁾。プログラム評価を検討するにあたり重要なのが、“当該プログラムはどのように定義できるか”，という点を明らかにすることである。これはまさに、プログラムという概念を曖昧にせず見極めることに他ならない。例えば、プログラム評価は政策評価を行う際の1つの原型(プロトタイプ)とされることがあり、そこでのプログラムには、①解決したい課題、②解決を促す政策的手段案の作成手順、③適切な手段の選定規準(後の評価規準に相応)、④政策デザイン(目標と手段の適切な組み合わせ)の運用マニュアル、⑤スケジュール、⑥必要な資源と入手先が明記されている¹⁰⁾。

プログラムは多種多様であり、介入方法(例：直接的介入・間接的介入)や介入レベル(例：個人・組織・地域コミュニティ)、そして評価主体(例：自己評価・外部評価)も様々である。いずれにせよ、プログラムが問題やニーズを抱えた利用者(あるいは対象となる地域や組織)に対して意図した付加価値をどの程度もたらすことができるかが、試金石となる。具体的には、プログラムの価値を効果(effectiveness)、成果(outcome)、インパクト(impact)といった形で、利用者側・実施者側の視点を踏まえ適切な評価デザインおよび規準(criteria)・基準(standard)を基に測定し、判断する行為がいわば狭義のプログラム評価ということになる。

一方で、広義のプログラム評価には様々な定義が存在する。例えば、社会科学アプローチによる

^{*1} Program Evaluation

^{*2} Tomoyuki Yasuda 西武文理大学

評価体系を構築したロッシ (P. H. Rossi) らは、「評価とは、社会調査の方法を活用し、社会プログラム等による介入の効果を体系的に研究することである。これらのプログラムは、政策的・組織的な文脈・環境において用いられるものであり、社会状況を改善するためのソーシャルアクションの情報源となるものである」としている¹¹⁾。また実用重視型評価 (utilization-focused evaluation) を提唱したパットン (M. Q. Patton) は、「プログラム評価とは、プログラムの活動、性質、アウトカムの情報を体系的に収集し、当該プログラムについて何らかの判断を下し、プログラム介入による効果の改善を行い、将来のプログラムについての決定を行うことである」としている⁶⁾ (p.23)。さらに、理論主導型評価 (theory-driven evaluation) を提唱したチェンは、「プログラム評価とは、プログラムを体系的に査定し、プログラムの計画、実施、効果を向上するために、評価のアプローチ、テクニック、知識を応用することである」¹²⁾としている。これらの定義からも、プログラム評価は、人々の意識・態度・行動などを明らかにしようとする通常の社会調査や実験データによって仮説検証を行う心理実験とは違い、ア

クション志向が強いアプローチであることが分かる^{*1)}。

2. プロセス評価とアウトカム評価

プログラム評価には様々なアプローチや方法論が存在するが、それらを包括的に捉える枠組みに「プロセス評価」と「アウトカム評価」がある。ここでは、方法論の開発が急がれる前者のプロセス評価に重点を置きつつ、これら2つの評価形式について概説する⁹⁾。

(1) プロセス評価

良い評価結果が出れば、当該プログラムは成功したと結論づけることができる。しかしこれでは、何故、どうやってその良い結果が得られたのかという疑問には答えられてはいない。この疑問点に答えるためには、プログラム実施の中間地点において介入メカニズムを体系的に査定し、実施状況の中間評価が必要となる。これがプロセス評価である。

プロセス評価は、より専門的には「対象となる集団に、意図されたとおりにサービスが届いているかの判断を行う評価」と定義されている¹³⁾。すなわちプロセス評価は、プログラムを“ブラッ

*1 またプログラム評価は「特定の目的を持って設計・実施されるさまざまなレベルの介入活動およびその機能についての体系的査定であり、その結果が当該介入活動や機能に価値を付与するとともに、後の意思決定に有用な情報を収集・提示することを目的として行われる包括的な探究活動」と定義することも出来る¹⁴⁾ (安田・渡辺, 2008, p. 5)。

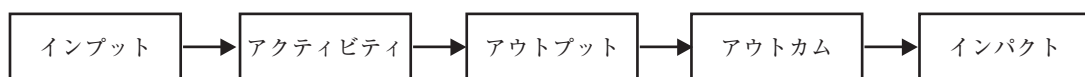
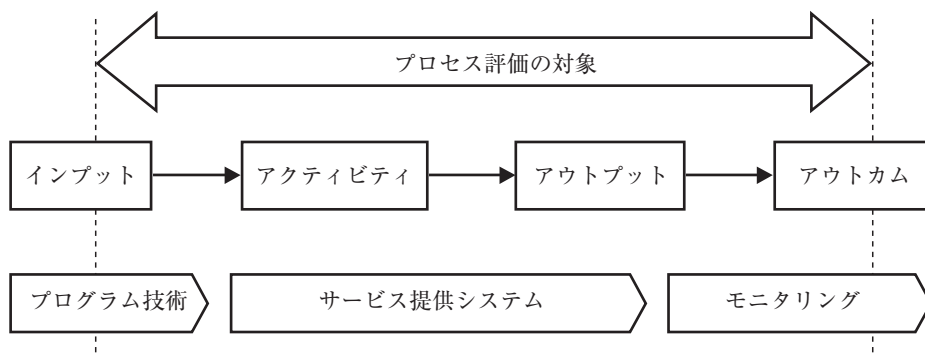
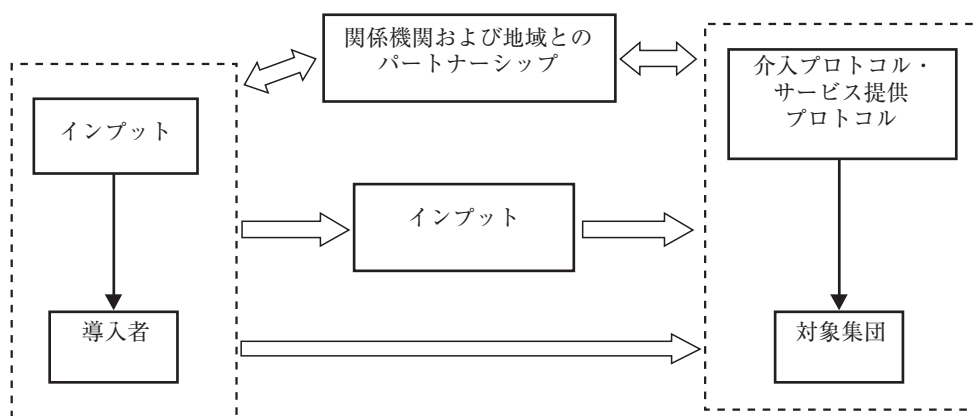
*2 「インプット」とは投入資源のことを指し、プログラムの運営者やサービス提供者、実施設備 (実施場所) や諸費用 (人件費・教材費・施設利用費)、関連情報の収集、組織との連絡調整といったプログラムに直接的・間接的に投入される資源 (ヒト・モノ・カネ・情報・時間・ネットワーク等) のことを指す。

*3 「アクティビティ」とはプログラムの目的やゴールを達成するために実際に行われる活動のことを指し、本研究で取り上げる健康増進プログラムにおいては主に、健康問題等の発生の予防 (prevention) や心身の健康や自己の促進 (promotion) を目標とする活動である。

*4 「アウトプット」とは、プログラムの実施によって直接的にもたらされた結果すなわち“出力”のことを指す。プログラムが実施されれば、その仕事量がデータ (例: 数値・記録) として現れてくる。例えば、参加者の数や参加率、参加者の特徴や属性などが具体的なアウトプットである。

*5 「アウトカム」とは、プログラムが参加者にもたらした“メリット (merit)”や“付加価値 (added value)”のことを意味し、主に、参加者に生じた知識・態度・スキル・興味関心などの変化・変容のことを指す。またコミュニティ・レベルでの成果もアウトカムと位置づけることができる (例: 地域における社会関係性が高まる)。

*6 「インパクト」とは、プログラム実施後に、派生的・副次的に生じる影響のことで、プログラムの参加者への直接的な効果を超え、自治体や地域社会に及ぼす影響 (例: コミュニティ活性化、“健康資本”の強化、医療費・社会保障費の抑制) などがそれにあたる¹⁶⁾。

図1 ロジックモデルの基本構造¹⁵⁾図2 プロセス評価の焦点とアプローチ⁹⁾ (p. 145)図3 アクションモデルの概要¹²⁾

クボックス”にしないための手段であると言える。その点を理解する上で有用な枠組みにロジックモデル (logic model) がある。ロジックモデルは、プログラムをインプット (投入資源: input)^{*2}, アクティビティ (活動: activity)^{*3}, アウトプット (結果: output)^{*4}, アウトカム (成果: outcome)^{*5}, インパクト (impact)^{*6} に分けて考え、各要因および要因間に存在する関係性を、“もし～ならば (if), …する (then)” という一定の関係性の連鎖によって可視化したものであり、基本構造は図1のように示される¹⁵⁾。

このロジックモデルを参考にすると、プロセス評価の焦点とアプローチは、図2に示された範囲として認識される⁹⁾。つまり、プログラムの投入

資源 (インプット) がどのように活動 (アクティビティ) に用いられ、その結果 (アウトプット), プログラムがどのように活用され (例: 利用状況・参加状況), 最後に当該時点において実際に利用者・参加者の効果・成果 (アウトカム) に結びついているか, という一連の流れを評価するのがプロセス評価となる。

例えば, 「プログラム技術」を評価するにあたっては, アクティビティ (サービスやプログラム) を提供するための“プログラミング”の方法, つまりプログラム実施に当たって必要となる資源 (ヒト・モノ・カネ・情報) の有無や多寡が評価される。また, プログラムの実行能力 (capacity) や実施・導入状況 (implementation) に焦点を当て

た評価が行われることも、包括的なプロセス評価において重要な視点であると言える¹⁷⁾。

また「サービス提供システム」の評価を検討する上では、前述のプログラム技術の評価と合わせて、プログラムに関するアクションモデル (action model) の枠組み¹²⁾が有用となる (図3)。これは、プログラムの対象集団にサービスが提供されるまでの道筋を示したものであり、特に、プログラムを実施するにあたっての他機関や地域とのパートナーシップや組織特性を含む様々な社会生態学的 (ecological) な文脈の影響が検討されているところに特徴がある。実際、地域の健康増進プログラムに関する調査研究では、実施主体の組織特性 (例：リーダーシップ、コミュニティ感覚) がプログラムの活動 (アクティビティ) を実施するにあたって肯定的な役割を果たすことが明らかになっている^{18,19)}。

さらにアウトプットからアウトカムの流れに関する「モニタリング」を行うにあたっては、プログラムやサービスの利用状況や提供状況などが検討される。具体的には、利用状況については、①活動内容、②提供期間・頻度、③利用者数と属性、④サービス提供者の特徴といった点を中心に²⁰⁾、提供状況については、①実施段取り、②サービスの程度 (“投与量”)、③実施上の問題点や否定的効果、④計画と実施の整合性、⑤安定性、⑥利用ニーズへの対応などが検討されることになる⁹⁾。

また、心理教育や人材育成等を目的とした研修プログラムのモニタリングを行う際には、参加者自身に“学びのプロセス (状況)”の評定を求めることも可能である。例えば、中高年者の職業観の再発見やメンタルヘルス向上などを通じた“人生再生”を目的とした研修プログラム²¹⁾では、受講生の「知識」 (例：新たな発見があった、すでに持っている知識と結びつけられた、自分の考えが整理できた)、「態度」 (例：真剣に取り組めた、自分の課題に向き合えた)、「興味・関心」 (例：講義内容に対して興味をもった、講義内容をもっと深く知りたい) に関するモニタリング (プロセス) 指標が設定され、講義毎のデータ収集が行われた。分析の結果、研修プログラムを受けるにあたっての知識の習得や態度や興味・関心

が高かった受講者に対して、より高い成果・効果 (例：「抑うつ」の向上) が得られることが明らかになった。

(2)アウトカム評価

“プログラムの効果・成果はあったのか”を査定・評価するのがアウトカム評価である。既述の通り、プロセス評価においても当該時点における参加者の効果・成果に関する評価を行うことがあるが、アウトカム評価では、プログラムの実施後、事後的に効果測定・評価が行われることになる。アウトカムは、主に利用者の①行動・行為、②意識・意欲、③認知・態度、④知識・理解、⑤興味・関心、⑥スキル・技術、⑦状態・立場などの“変化”として現れるため、それらの特定・指標化を行い、プログラムの効果が把握されることになる⁶⁾。アウトカム評価の代表的形式に、実験・準実験デザインがあるが、紙幅の関係上、ここでは省略する。詳しくは、文献14)あるいは文献9) (p.192-228)を参照されたい。

3. 評価者の役割

プログラム評価における評価者の役割は、当然のことながら、評価を行うことである。より専門的な役割を考えていく上では、理論家 (i.e., プログラムの実施・運営に関する理論的背景を構築する個人やグループ) および実践家 (i.e., プログラム介入を行うスタッフ) との関係性に着目する必要がある²²⁾。

まず理論家に対しては、“実践現場においてプログラムの理論は通用するのか”という点を問う姿勢が評価者として重要となる。つまり、プログラムは理論的に正しく機能しそうであるか、現実とはかけ離れた仮定や理想論は含まれていないか、ということを事実や理論家とのコミュニケーションを通して明らかにすることである。

一方、実践家に対しては、“行っている実践を理論的に説明するとどうなるのか”ということを問う姿勢が大切となり、ここでは、現場ではうまくいっている実践であるが、その流れや利用者への効果を理論的 (機能・構造的) に説明するとどうなるのかを突きつめていく作業が伴うことになる。一般に、プログラム評価には、①アカウンタ

ビリティのための評価, ②プログラムの改善・質向上のための評価, ③価値判断のための評価, ④評価研究のための評価という4つの目的がある。評価者の役割は評価目的別に異なるが⁶⁾, これらをまとめたものが表1である。

4. アカウンタビリティ向上：より現実的な評価の選択肢として

エビデンスに基づいた実践 (evidence-based practice) は、「研究による最適な科学的根拠に臨床の専門的知識・技術と患者の価値尊重を統合したもの」²³⁾と定義されており、特に心理学領域では、「研究による最適な科学的根拠にクライアントの特徴・文化・嗜好を考慮した上での臨床の専門的知識・技術を統合したもの」²⁴⁾とされている。具体的に、ここでの“研究による最適な科学的根拠 (best research evidence)”とは、ランダム割付か否か、比較群が存在するか、サンプル数は十分であるか、統計的有意性があると同時に効果量の程度は高いか、といった方法論の厳格性 (methodological rigor) に拠るところが大きい。

一方、ランダム化比較試験 (RCT: randomized control trial) に代表される方法論の厳格性を担保するための実験デザインを用いたプログラム評価は、対人援助・コミュニティ援助を中心としたヒューマンサービス領域では未だ数少ないのが現状である。利用者やクライアントへの倫理的配慮や現場の理解 (“buy-in”), そして現場への方法論の技術移転 (technology transfer) が十分でない現状を踏まえると、実験デザインのような厳格な評価方法を用いた評価に耐えうるプログラムはごくわずかと予想される¹⁶⁾。このようななか、“サービス提供”と“エビデンス提示”との狭間

に存在する課題を解決する方法としてアカウンタビリティの向上が考えられる。

アカウンタビリティとは説明責任を意味し、米国における1993年の政府業績成果法 (GPRA: Government Performance and Results Act), および日本の政策評価法 (2004年) の制定以降、社会・公共サービスの評価において特にその役割が注目されてきた。またこれらと連動して、臨床心理や社会福祉などのヒューマンサービスの領域においても、アカウンタビリティの遂行が意識され始めたという経緯がある。

具体的に、プログラム評価におけるアカウンタビリティ向上の方策を考えるにあたっては、まずプログラムの活動方針やゴールを明確にすることが肝要であると言える (ステップⅠ)。すでに実施されているプログラムであれば、その方向性を再確認・再検討する意味でも、活動方針やゴールの明確化は意義があると言える。

次に、プログラムやサービスを実施した結果、“利用者・受益者にどのような変化が起こるか (起こって欲しいか)”という点をインパクト理論によって可視化することも可能であろう (ステップⅡ)。そして、前述のロジックモデルにより、実際にそのような変化を起こすためのプログラムの実施・運営の可視化が有用である (ステップⅢ)。

これらの情報を参考にしつつ、ロジックモデルの枠組みを援用して、実際に評価を行うにあたって検討する (すべき) 点を評価クエスチョンとして抽出し (ステップⅣ)、データ収集・分析を行う (ステップⅤ)。以上の5つのステップは、プログラムを体系的に捉え、アカウンタビリティ向上を目指す上で有効な方法論と言えよう¹⁶⁾。

表1 評価目的別にみた評価者の役割とスキル⁹⁾

評価目的	期待される役割	必要とされるスキルや資質
1. アカウンタビリティ	査定・聞き取り役	綿密性・完全性に基づく情報処理能力
2. プログラムの改善・質向上	コンサルタント役	問題点を見抜く洞察力や現場とのコラボレーション能力
3. 価値判断・意思決定	裁判官・審判役	権威的な立場からの中立性と決断力
4. 評価研究	研究者役	評価に関する理論・方法論の構築スキル

プログラムの状況が実践者にとっても“見えにくい”場合、あるいはより抽象度が高い取り組みも存在する。各実践者がそれぞれの援助・サービス提供を行っているものの、それらを“プログラム”として捉えるには個別具体例が多すぎる取り組みである（例：臨床心理学領域における臨床心理地域援助）。これらをプログラム評価の枠組みのなかで理解するにあたって有用なのが、ステークホルダー情報に基づいたプログラムの構造化である²⁵⁾。1つのアプローチとしては、ステークホルダーに対して簡易的な情報収集を行い、その結果をロジックモデルの枠組みを用いて分類する。そしてその分類された記述をもとに、ベンチマークを作成し、援助が行われている各フィールドでの評価に用いるという視座である。このアプローチは、例えば、プログラムのベストプラクティス（best practice）やそれに沿ったガイドラインを整備する上でも有用であると考え²⁵⁾。

5. 評価キャパシティ構築

今後のプログラム評価研究を展開する上で、評価キャパシティの構築も重要となってくる。キャパシティとはすなわち実行能力のことであり、その有無や多寡は、プログラムの成否に直接的・間接的な影響を及ぼすことになる。そのため、プログラムやそれを提供する組織の実行能力を開発し、それを維持・促進するための方策としてキャパシティ構築（capacity building）が重要となる。

特に、“評価を通したキャパシティ構築”を意味する「評価キャパシティ構築」（ECB: evaluation capacity building）の役割に期待が集まっている^{26, 27)}。具体的にECBは、「コンテキスト（組織的・社会的な文脈）に依拠する意図的なアクションのシステムで、そのシステムによるプロセスや実践が、質の高いプログラム評価およびそれを適切に活用することが日常のかつ継続的な実践となる状態を、1つまたはそれ以上の組織、プログラム、現場において創出・維持することである」²⁸⁾と定義されている。端的に言えば、評価のプロセスを日々の実践の質向上、健全な組織文化の涵養に役立てることが、ECBにつながるという訳である。

ECBが果たす役割への期待は、様々なレベルに現れる効用にあると言ってよい。例えば、個人レベルの効用としては、プログラムの実践者やサービスの提供者の認知・行動・情動の変化、すなわち評価を行うための知識やスキルが習得され、評価思考（evaluation thinking）が普段の実践に応用され、評価に対して肯定的態度がとられるようになり、評価活動への積極的な関与につながる、という報告がある²⁹⁾。また組織レベルへの効用としては、組織学習（organizational learning）が推進され、評価結果が継続的に活用されることにより、パフォーマンスが向上するという点にあるとされている²⁹⁾。

またECBの理解には、①「外部環境」（例：当該組織に対するアカウンタビリティの要請）、②「組織構造」（例：組織の評価者としての役割の明確化）、③「組織文化」（例：組織の慣習、価値、前提の共有と新たな施策への批判的吟味）、④「組織リーダーシップ」（例：評価に関するマネジメント面での支援）の少なくとも4つの組織要因との関連性が検討される必要があるとされている²⁶⁾。

6. 結 語

本セッションでは、プログラム評価の概要およびプログラム評価研究の動向について概説してきた。プログラム評価には、ある一定の決まった方法論は存在する訳ではない。したがって、従来の実験研究法、統計・調査手法、インタビューやフォーカスグループなどの定性的方法など、プログラム評価に応用できる方法論は数多い。今後の方向性としては、これらの方法論を駆使したプログラム評価、特にプロセス評価の方法論の開発・発展が望まれるとともに、アカウンタビリティ向上を推進するための方策や評価キャパシティの構築に向けたフィールドでの技術移転のあり方の検討が必要であろう。

■文 献

- 1) 大西弘高. 新医学教育学入門. 教育者中心から学習者中心へ, 医学書院, 東京, 2005.

- 2) Campbell DT & Stanley JC. Experimental and Quasi-Experimental Designs for Research. Skokie, IL: Rand McNally. 1963.
- 3) Cook TD & Campbell DT. Quasi-Experimentation: Design & Analysis Issues for Field Settings. Houghton Mifflin Company, Boston, 1979.
- 4) Cambell DT. Reform as Experiments. *American Psychologist* 1969; **24**: 409-29.
- 5) Shadish WR & Luellen JK. Donald Campbell. The accidental evaluator. In M. C. Alkin (ed) Evaluation roots. Tracing theorists' views and influences (2nd edition). Sage publication, 2013. p.61-5.
- 6) Patton MQ. Utilization-focused evaluation (3rd ed.) Sage publication. 1997.
- 7) Fetterman DM, Kaftarian SJ & Wandersman A. Empowerment Evaluation. Knowledge and tools for self-assessment and accountability. Sage Publication, 1996.
- 8) Chen HT. Theory-Driven Evaluations. Sage publication, 1990.
- 9) 安田節之. プログラム評価, 対人・コミュニティ援助の質を高めるために. 新曜社, 東京, 2011.
- 10) 山谷清志. 政策評価. ミネルヴァ書房, 京都, 2011, 193.
- 11) Rossi PH, Freeman HF & Lipsey MW. Evaluation. A Systematic Approach (6th ed.). Thousand Oaks, CA: Sage. 1999, p.4.
- 12) Chen HT. Practical program evaluation: Assessing and improving planning, implementation, and effectiveness. Sage publication, 2005, p.3, p.29.
- 13) Scriven M. Evaluation Thesaurus (4th ed.). Newbury Park, CA. Sage publication. 1991.
- 14) 安田節之, 渡辺直登. プログラム評価研究の方法. 新曜社, 東京, 2008, p.119-49.
- 15) W. K. Kellogg Foundation. Using logic models to bring together planning, evaluation, & action: Logic model development guide. Battle Creek, MI: Author, 2001, p.1.
- 16) 安田節之. プログラム評価研究, 臨床心理サービスのアカウントビリティ向上に役立つ視点. 臨床心理学 2013 ; **13** : 337-42.
- 17) 安田節之. 地域を基盤としたプログラムにおける組織プロセスが評価キャパシティ構築と評価有益性意識に及ぼす影響. 西武文理大学サービス経営学部紀要 2013 ; **22** : 13-28.
- 18) 安田節之. コミュニティ介入の効果を高める組織特性の検討, 臨床心理地域援助における評価研究の試みとして. 臨床心理学, 印刷中.
- 19) 安田節之. ロジックモデルを規準とした測定指標によるコミュニティ介入のアセスメント, 地域リーダー研修の調査データより, コミュニティ心理学研究, 投稿中 a.
- 20) Weiss, C. Evaluation. Methods for studying programs and policies (2nd ed.). Prentice Hall: Upper Saddle River, NJ.1998.
- 21) 牧野篤, 安田節之, 人生リバイバルプログラム実施委員会. つながりが人を救う, 人生リバイバルプログラム実施報告書. 東京大学大学院教育学研究科学習基盤社会研究・調査モノグラフ. 2012.
- 22) Weiss C. Nothing as practical as a good theory: exploring theory-based evaluation for comprehensive community initiatives for families and children. In. J. Connell, A. Kubisch, L. Schorr and C. Weiss. (eds.) New Approaches to Evaluating Community Initiatiaves. Aspen Institute.1995.
- 23) Institute of Medicine. Crossing the quality chasm: A new health system for the 21st century. National Academy Press, Washington, DC, 2001, 147.
- 24) APA Presidential Task Force on Evidence-Based Practice. *Evidence-based practice in psychology* 2006; **61**: 271-85.
- 25) 安田節之. ステークホルダー情報に基づいた予防介入プログラムの構造化, ベンチマーキングによる評価研究の予備調査として. 西武文理大学サービス経営学部紀要, 投稿中.
- 26) Bourgeois I & Cousins JB. Understanding dimensions of organizational evaluation capacity. *American Journal of Evaluation* 2013; **34**: 299-319.
- 27) Compton DW, Baizerman M & Stockdill SH. The art, craft, and science of evaluation capacity building. *New Direction of Evaluation* 2002; **93**.
- 28) Stockdill SH, Baizerman M & Compton DW. Toward a definition of the ECB process. A conversation with the ECB literature. In Compton DW, Baizerman M & Stockdill S (Eds.), The art, craft and science of evaluation capacity building.

- New Directions for Evaluation* 2002; **93**: 7-25.
- 29) Garcia-Iriarte E, Suarez-Balcazar Y, Taylor-Ritzler T & Luna M. A. catalyst-for-change approach to evaluation capacity building. *American Journal of evaluation* 2011; **32**: 168-82.
- 30) Cousins, J. B., Aubry, T., Smith-Fowler, H., & Smith, M. (2004). Using key component profiles for the evaluation of program implementation in intensive mental health case management. *Evaluation and Program Planning* **27**; 1-23.