

発達研究における縦断的アプローチの役割と方法論：What, Why and How?

伊藤 大幸

(お茶の水女子大学基幹研究院人間科学系)

縦断研究は、対象の経時的な変化の様相と機序を解明するための科学的手法として確立された地位を築いているが、発達研究においてどのような強みを持ち、また、どのような限界や陥穽を有するのかについて、必ずしも明確な認識が共有されていない。そこで本稿では、実際の発達研究の文脈に即して、縦断研究とは何か (What)、なぜ必要なのか (Why)、どのように進めるか (How) について多面的に議論した。縦断研究には、測定方法、測定間隔・期間、要因操作の有無などの点で異なる多様な研究デザインが含まれるが、これらに共通しているのは、同一の対象から複数時点でのデータを収集し、個人内の時系列的な変化や関連に着目した分析を行う点である。こうした特徴を持つ縦断的アプローチが発達研究において発揮する有効性は、(1) 年齢・コホート・時期の効果を分離できること、(2) 個人内変動の軌跡とその個人差を定量化できること、(3) 因果関係 (特に時間的順序性) に関する手がかりを得られることの3点に整理できる。縦断的アプローチにより、個人内の時間的変動という軸が加わることで、現象を捉える視点は飛躍的に複雑さを増す。こうした視点の複雑さは、研究デザインやデータ分析の方法論の複雑化という代償とともに、横断研究とは全く異なる種類の創造的なりサーチクエスションの創出をもたらすことで、発達研究の広がりや深まりに貢献するであろう。

【キーワード】 縦断研究, 年齢—時期—コホート分析, 発達軌跡, 因果推論

縦断研究は、対象の経時的な変化の様相と機序を解明するための科学的手法として、心理学に限らず、医学、生物学、経済学など、幅広い分野において確立された地位を築いている。とりわけ個体発達の過程を扱う発達心理学において、縦断研究は中核的な研究手法として認識されており、この手法が発達の解明に不可欠な貢献を果たすという点に関しては、おおむね共通認識が形成されていると言える。

しかし、そもそも縦断研究とはどのような研究を指すのか (What)、縦断研究がなぜ重要なのか (Why)、縦断研究をどのように進めるべきなのか (How) については、発達研究者の間でも明確な認識が共有されていないように思われる。例えば、横断研究の論文には、「研究の限界として「今後、縦断研究による検証が必要である」といった記載がなされることが多いが、なぜ、どのような縦断研究が必要なのか、明確に記述されることはほとんどない。また、縦断的手法を用いた研究でも、2変数を経時的に測定し、交差遅延モデルにより双方向の効果を推定するだけで、因果関係が証明されたかのような結論を導いている研究が少なくない。漠然とした縦断研究のイメージが独り歩きし、リサーチクエスションや方法を問わず縦断研究が無条件に優れたエビデンスをもたらすかのような誇張された認識が広まっていることは、発達研究の進展にとって望ましい状況とは言えない。

そこで本稿では、実際の発達研究の文脈に即して、縦断研究の What, Why, How について議論することを通じて、縦断研究がどのような強みを持ち、また、どのような限界や陥穽を有するのか、できる限り具体的に論じていきたい。本稿の目的は、縦断研究を礼賛することでも否定することでもなく、縦断研究の特徴と方法論を批判的に論じることで、縦断研究がより効果的に発達の解明に貢献できる状況の形成に寄与することである。

児童・青年の情緒・行動問題予防のための大規模縦断研究

本稿では、縦断研究の役割と方法論を論じる上で、著者が実施している大規模縦断研究のデータを事例として取り上げる。この研究は、主に思春期に顕在化する不登校、いじめ、自傷行為、非行などの情緒・行動上の問題の発生機序を検証し、問題の予防の方策を見出すことを目的としている。中部地域の中規模都市との間で研究と支援に関する協定を締結し、2008年から現在まで継続して、市内の全ての保育所・小中学校での縦断調査を1年間隔で実施している。調査結果のフィードバック、児童生徒向けの心理教育、教員向けの研修会、保護者向けの講演会、適応指導教室でのカウンセリングなど、研究知見に基づく多様な支援活動を並行実施することで市・学校園・家庭との協力関係を構築し、9割を超える

有効回答率を維持している（約1万2千人）。これまでの14年間の調査で、25の出生年コホートの計2万4千人余りから、のべ約16万（人×時点）のデータが得られている。発達心理学領域における縦断研究としては、調査期間・人数の両面で国内最大規模の研究である。

調査内容としては、結果変数である内在化問題（抑うつ・不安、不登校、自傷行為、摂食行動異常など）および外在化問題（攻撃性、いじめ加害、非行、触法行為など）に加え、これらに影響を及ぼすと考えられる個人要因（気質、知能、発達障害特性、学業成績、性別違和感など）および環境要因（社会経済的状態、養育行動、友人関係、教師の指導スタイルなど）を幅広く調査している。測定には主に質問紙を用いているが、評定バイアスに対処するため、保護者、教師（保育士）、児童生徒の三者から評定を得ている。本研究の成果は、これまで60篇以上の論文により報告されている（e.g. 伊藤ほか, 2017, 2021; Murayama et al., 2021; Katagiri et al., 2021）。

縦断研究とは何か

発達心理学領域において、「縦断研究」という用語から連想される典型的な研究は、特定の集団に対して、数か月から数年の間隔で、複数回の追跡的な調査（主に質問紙調査）を行った研究であろう。上述の著者らの縦断研究もこれに該当する。しかし、これは縦断研究の典型例ではあっても定義ではない。縦断研究の定義には様々なレベルのものがあるが、最も広い定義では、特定の対象について複数の時点におけるデータを収集し、経時的な変化を検証した研究全般を指している。この定義に基づけば、心理学領域の非常に幅広い範囲の研究が縦断研究の範疇に含まれると言える。こうした多様な縦断研究を特徴づける次元として、少なくとも以下の6つがある。

①データ収集の時間的な方向性。心理学領域の縦断研究は、特定の対象をある時点から一定期間にわたって追跡して調査を行う前方視的研究が大部分を占めるが、医学（疫学）領域では、何らかの疾患に罹患した人々について、成育歴、既往歴、生活歴などを遡って調査し、疾患のリスク要因を探るといった後方視的研究も多く行われている（e.g., Yao et al., 2020）。後方視的研究は、収集される情報の正確性という点で前方視的研究に劣る場合が多いが、実施コストが低いこと、サンプルの脱落が生じないこと、仮説的な原因—結果の時間差に関する情報が得られることなどのメリットもある。

②測定の方法。質問紙調査だけでなく、行動観察、認知課題、生理学的指標などを用いた研究も縦断研究に含まれる。また、質問紙調査も、自己報告だけでなく、保護者、教師、友人などの他者による報告を含む場合がある。

③要因の操作や介入の有無。縦断研究には、要因の操作を行わない相関研究だけでなく、要因の操作を行う実験研究や治療・支援などを行う介入研究も含まれる。例えば、心理療法プログラムの効果検証において、プログラムの開始前、実施中、終了直後に加え、数カ月や数年にわたってフォローアップ調査を行い、効果の持続性を検証するといった研究も縦断研究の一種である。

④測定の間隔および期間。数年間隔で50年にわたって調査を行うような長期縦断研究もあれば、日誌法や経験サンプリングなどの手法を用いて、1日あるいは数時間の間隔で1週間の調査を行うといったインテンシブな縦断研究（本特集の児玉・榊原（2022）など）もある。また、突き詰めて考えれば、単一の実験内での行動や生理学的指標の反復測定も、その経時的な変化に着目するのであれば一種の縦断研究として位置づけられる。

⑤各時点の測定内容の異同。典型的には、各時点で同一の変数セットについて測定を行い、各変数の変化や相互の関連を検証することが多いが、長期縦断研究などでは、発達段階によって経験する事象に差異があるため、時点によって異なる内容を測定することもある（例えば、学齢期のいじめ、成人期の就職、転職など）。

⑥サンプルサイズ。集団を対象とした研究だけでなく、単一の個人を対象とした研究も縦断研究に含まれる。例えば、応用行動分析の領域で広く用いられる単一事例実験（n-of-1 試験）では、単一（または少数）の個人について、介入を実施している期間と実施していない期間の間の変化に基づいて介入の効果を検証する（詳細は本特集の山田（2022）の寄稿を参照）。こうした研究も個人内の経時的変化を扱っている点で縦断研究に該当する。

以上のように、縦断研究にはきわめて多様な研究が含まれるが、これらに唯一共通しているのは、同一の対象から複数時点でのデータを収集し、個人内（within-person）の時系列的な変化や関連に着目した分析を行うという点である。これは、一時点でのデータのみが収集され、個人間（between-person）の差や関連だけが分析の対象となる横断研究とは明確に区別できる特徴である。

こうした特徴を持つ縦断的アプローチが発達研究において発揮する有効性は、大きく以下の3点に整理できる。すなわち、(1) 年齢・コホート・時期の効果を分離できること、(2) 個人内変動の軌跡とその個人差を定量化できること、(3) 因果関係に関する手がかりを得られることの3点である。発達心理学領域において縦断研究が特別な地位を与えられているのは、こうした方法論的な強みが個体発達のプロセスとメカニズムを明らかにする上で有利または不可欠な要件として機能するためである。以下、この3つの有効性について、著者らの研究

データを例として取り上げながら、具体的に論じていく。

年齢・時期・コホートの効果の分離

発達研究の最も基本的な目的は、年齢にともなう標準的 (normative) な個体発達の様相を記述することである。Erikson (1959) や Piaget (1970) の古典的な発達段階理論を例に挙げるまでもなく、人の認知・感情・行動の標準的な発達の過程を記述することは、発達研究の第一義的な目的である。母子手帳に掲載された身長・体重の発育曲線のように、標準的な発達の様相が明らかになることで初めて、それを基準として個々人の発達の状態を評価し、必要に応じて適切な支援につなげることが可能になる。

標準的な発達を記述するにあたり、一般的には、縦断研究よりも横断研究が用いられることが多い。例えば、発達検査や知能検査を開発する際には、通常、一時点で幅広い年齢層のデータを収集した上で、個々の年齢層における得点分布に基づいて基準値 (標準得点の換算表) が算出される。こうした方法は、検査開発のコストを抑える簡便法としては合理的であるが、より普遍的な年齢の効果を検証しようという発達研究の文脈では、必ずしも適切な方法とは言えない。なぜなら、一時点の横断データにおいては、年齢による得点分布の違いが、年齢自体の効果を反映しているのか、「コホート」(いつ出生したか) の効果を反映しているのか、区別することができないためである。言い方を変えれば、単独の横断データ上では年齢とコホートの効果が完全に交絡しており (年齢とコホートの相関係数が1となる)、両者を分離して推定することができない。また、一時点のデータであるため、時間に関連するもう一つの変数である「時期」(いつ測定が行われたか) の効果を推定することもできない。

この問題に対して、縦断研究が無条件に解決策を提供するわけではなく、デザインと解析の両面において周到な工夫が必要となる。デザインに関して言えば、単一の出生年コホートを対象とした縦断研究では、コホートが統制されているものの、年齢と時期の交絡が生じるため、やはり両者の効果を分離することができず、統制されているコホートの効果も推定できない。つまり、一時点の横断研究と状況は大きく変わらない。年齢、コホート、時期の効果を分離して評価するためには、複数のコホートに対して、複数の時点で調査を行うことが必要条件となる。

解析においても、いくつかの困難な課題に対処する必要がある。第一に、縦断的に測定されたデータは、複数時点の測定値が各個人にネストされた階層構造を成している。この場合、同一個人内の測定値は互いに相関を有

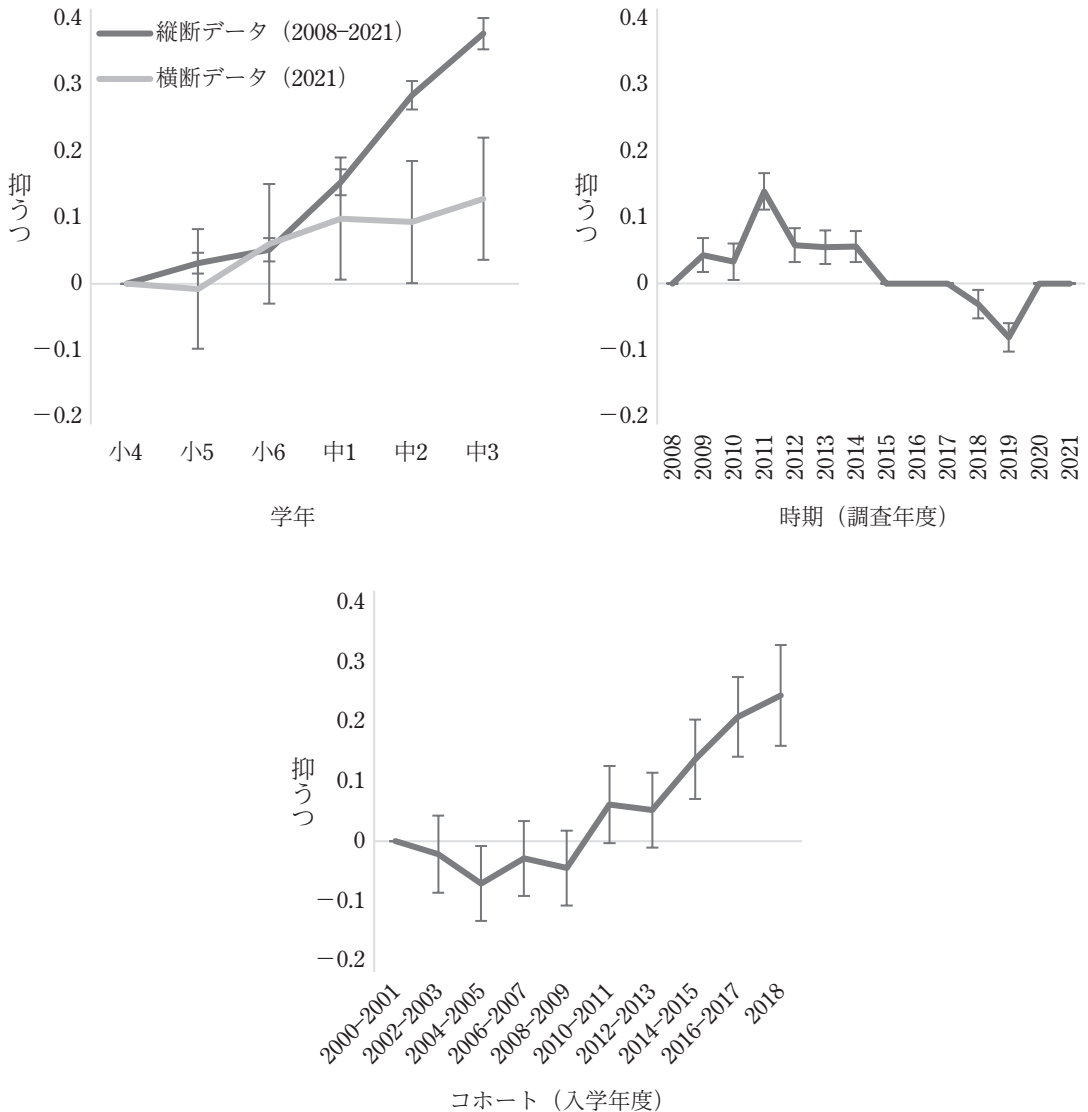
し、データの独立性が保たれないため、通常の線形回帰分析では正確な結果を得ることができない。第二に、年齢、時期、コホートの3変数の間には「年齢=時期-コホート」という関係が成立している。これは、いずれか2つの変数によって、残り1つの変数の値が一意に定まる線形従属という関係に該当する。こうした関係が成立している場合、3つの変数をそのままの形で説明変数として投入すれば、多重共線性により解を求めることができない識別不定の状態に陥る。

抑うつに対する学年・時期・コホートの効果の推定

Figure 1 は著者らの縦断研究のデータを用いて、抑うつに対する年齢 (学年)、時期 (調査年度)、コホート (入学年度) の効果を推定した結果である。17,388 名から得られたのべ 71,228 のデータを分析に使用している。ここでは上記の第一の問題 (データの階層性) に対処するため、測定時点をレベル1、個人をレベル2として、マルチレベルモデルによる推定を行った。3つの要因のうち、測定時点ごとに変動する年齢および時期はレベル1の変数として、時点ごとに変動せず、個人間でのみ変動を持つコホートはレベル2の変数として投入した。

第二の問題 (多重共線性) に関しては、万能の解決策は見出されておらず、異なる特徴を持った複数の方法が提唱されているが (Mason, Mason, Winsborough, & Poole, 1973; Yang & Land, 2013; Bell & Jones, 2015)、今回は古典的な Mason et al. (1973) の方法を拡張して用いた。Mason et al. (1973) の方法では、識別の問題を回避するため、3要因のいずれかについて、隣接する複数のカテゴリを併合する処理を行う。これは隣接するカテゴリ間で平均値が等しいという制約を課すことを意味する。ここでは、最もカテゴリ数が多く、隣接するカテゴリ間の抑うつの変動が小さかったコホートについて、隣接する2カテゴリずつを併合した。これにより、ひとまず線形従属は解消され、解を求めることが可能になるが、依然として要因間の相関は非常に強く、解の推定が不安定になる多重共線性は解消されない (Glenn, 2005)。著者らのデータでも多重共線性の指標となる VIF (10 を基準とすることが多い) は最も大きい項で 30 を超えた。そこで、モデルにさらなる制約を課すため、時期の要因について、各カテゴリのダミー変数をステップワイズ法の要領で順次投入し、有意な効果を示さなかったカテゴリの効果を0に固定した。その結果、Figure 1 のように、当初の基準カテゴリである 2008 年に加え、2015-2017 年および 2020-2021 年の各カテゴリの効果が0に固定された。これにより、VIF は最も大きいもので5程度にまで低減し、多重共線性はほぼ解消されたと判断できる。

学年の効果 (Figure 1 の左上図の黒のグラフ) については、小4から小6にかけて年に 0.03 SD 程度の緩やか



注. エラーバーは95%信頼区間を示す。抑うつは標準化得点を分析に使用。学年は小4、時期は2008、2015-2017、2020-2021、コホートは2000-2001を基準カテゴリとした。

Figure 1 マルチレベルモデルによる抑うつに対する学年・時期・コホートの効果の推定値

な傾きで上昇した後、小6から中3にかけては年0.10 SD程度より大きな傾きに切り替わっていることが見て取れる。この変化の様相は思春期にかけて抑うつが顕在化するという先行研究の知見 (McGuire, McCormick, Koch & Mendle, 2019) に整合的である。時期の効果 (Figure 1の右上図) については、2009年ごろにわずかな上昇が見られた後、2011年で再度の上昇が見られ、その後は2019年度まで低下傾向にあるものの、2020年

で三たびの上昇が生じている。時代背景と照合すると、2009年の上昇はリーマンショック後の経済状況の悪化、2011年の上昇は東日本大震災、2020年の上昇はコロナ禍の影響をそれぞれ反映している可能性が考えられる。コホートの効果 (Figure 1の下図) については、小学校入学が2004年度から2009年度ごろのコホートが最も低く、その後はおおむね直線的な上昇が見られる。ここ10年ほどの間に、児童生徒の抑うつの水準が(世代差

として)高まっていると言える。原因は定かでないが、何らかの時代的要因によって生じた生育環境の変化(例えば少子化、ネット利用の低年齢化、余暇活動の変化など)が抑うつへの脆弱性を高める方向に働いている可能性がある。

比較のために、2021年度の横断データに基づく学年の効果の推定値も示した(Figure 1の左上図の灰色のグラフ)。縦断データによる推定値と比較すると、学年による上昇の幅が小さいことが見て取れる。これは以下の通り、コホートの効果との交絡によって生じた結果の違いとして解釈できる。2021年度に小4であったコホートは2018年度入学、小5-小6であったコホートは2016-2017年度入学、中1-中2であったコホートは2014-2015年度入学、中3であったコホートは2012-2013年度入学にそれぞれ対応する。この4つのコホートでは、世代が後になるほど(単年度では学年が下がるほど)直線的に抑うつ水準が上昇しているため、その効果が調整されなかった場合、学年にともなう抑うつの上昇は過小推定されることになる。本研究のように、5,000名程度の大規模データにおいても、横断データと縦断データの間でこれだけの明確な結果の違いが生じたことは、注目に値する。サンプルサイズの拡大は、サンプル変動によるランダムな推定誤差を低減する上で有効な方法であるが、要因の交絡による系統的な誤差を低減する手段にはなりえない。

以上のように、コホート・時期との交絡を調整することで、年齢にともなう発達の変化について、より正確で普遍性の高い結論を得ることができる。これを可能にする複数コホート縦断研究は、発達心理学の中心的な関心変数である年齢の効果の推定する上で有効性の高い研究デザインと言える。また、年齢・時期・コホートの効果を分離して評価することは、個体側の成熟だけではなく、社会的な時代背景を含めた多面的な考察を可能にする。このことは生物学的要因と社会的要因の相互作用が織りなす個体発達の機序を明らかにするという発達研究のミッションを達成する上で重要な意味を持つ。

ただし、Yang & Land (2006)は、時点間で同一個人の紐づけを行わない反復横断データから年齢・時期・コホートの効果を推定するモデルを提唱している。この方法は不正確な推定値をもたらす可能性があることも指摘されているが(Bell & Jones, 2015)、少なくとも集団レベルでの標準的な発達の变化の様相を知る上で、縦断研究は(有効ではあるが)必ずしも必須のデザインではない可能性がある。

個人内変動の軌跡とその個人差の定量化

前項で扱ったような標準的な発達の様相を知ることが、発達研究の出発点として重要な課題であるが、発達

についてさらに理解を深めるには、各個人のレベルでの時系列的変化の軌跡(trajjectory)とその個人差を定量化することが必要となる。反復横断研究が代替手段となりえた集団レベルの標準的な変化の検証とは異なり、個人レベルの変化の軌跡の検証には時点間の同一個人の紐づけをとらざるを得ない縦断研究が不可欠となる。つまり、(1)変化の軌跡にどのような個人差があるか(例:道徳性の発達の軌跡にはどのようなパターンがあるか)、(2)変化の軌跡を何が予測するか(例:小児期逆境体験が生涯にわたる精神的健康の軌跡を予測するか)、(3)変化の軌跡によって何が予測されるか(例:情緒的症状の軌跡によって問題行動の発生が予測されるか)、(4)複数の変数の変化の軌跡に関連があるか(例:実行機能と心の理論の発達の变化は連動するか)、といった発達心理学上の重要な問いについて、縦断研究以外の手段で直接的に検証を行うことは原理的に困難である。これが発達研究における縦断的アプローチの最大の強みであると言える。

再び著者らの大規模縦断研究におけるデータを取り上げ、上記の(1)や(3)の問いに関する検証の例を示す。思春期に顕在化する深刻な内在化問題の一つとして不登校がある。不登校は一度生じると慢性化しやすく、卒業後の適応にも長期的な影響を及ぼすリスクがあることから(Cattan, Kamhöfer, Karlsson, & Nilsson, 2022)、未然に兆候を捉え、予防を図ることが重要である。著者らの調査において不登校(1学期段階で10回以上の欠席)を示した児童生徒の割合は、小4で0.89%、小5で1.01%、小6で1.32%、中1で1.66%、中2で3.73%、中3で4.44%と、学年とともに上昇していたが、特に中1から中2にかけて2%以上の顕著な上昇が見られた。そこで、この中2時点での不登校の発生を、前年度(中1)までの心理社会的適応の軌跡によって、どの程度の精度で予測しうるかを検証してみたい。

しかし、不登校を予測すると言っても、全ての生徒が同様のプロセスを経て不登校状態に至るとは限らない。ある生徒は学業不振、他のある生徒は友人関係の問題、さらに別の生徒は抑うつ症状を経て不登校状態になるかもしれない。こうした可能性を考慮すれば、全ての生徒に同一のモデルをあてはめて不登校を予測することは必ずしも妥当でない。そこで、ここではまず成長混合分布モデル(Growth Mixture Model; Muthen & Shedden, 1999)という手法を用いて、不登校状態に至る心理社会的適応の軌跡を類型化することを試みる。

不登校生徒の軌跡の類型

成長混合分布モデルは、変数の変化の軌跡を定量化する成長曲線モデルと、データから複数の未知の下位集団を同定する混合分布モデルを組み合わせた手法である(これらのモデルの詳細については本特集の西村(2022)

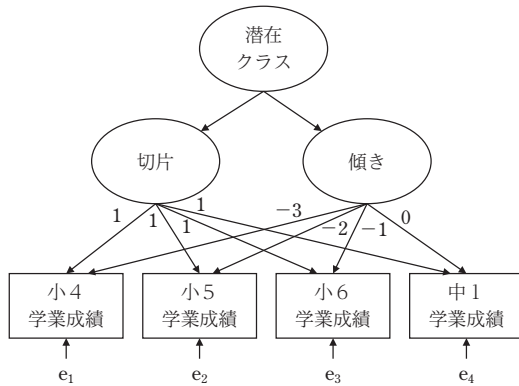


Figure 2 成長混合分布モデル

の寄稿を参照されたい)。一例として Figure 2 に学業成績に関する成長混合分布モデルを示す。図の下半分が成長曲線モデル，上半分が混合分布モデルに対応する。成長曲線モデルでは，複数時点で測定された変数について，その変化の軌跡を表す切片や傾きなどの成長因子を設定する。ここでは小4から中1の4時点での観測値の軌跡を，切片と一次の（直線的な）傾きによって表現するモデルをあてはめた。切片因子は全ての時点の観測値に対して1の負荷量を持った因子として指定される。一方，傾き因子は，直線的な変化を表現するために，時点を追うごとに負荷量（時間パラメータとも呼ぶ）が1ずつ変化する形で指定される。傾き因子の負荷量が0になる時点では，傾き因子が負荷せず，切片因子のみが負荷する形になるため，モデル上，切片はその時点における変数の真の値を表すことになる。今回は中2時点の不登校の発生を予測することが目的であり，直近の中1時点での各説明変数の水準が最も重要な意味を持つため，切片が中1時点の水準を表すように，中1における時間パラメータを0とし，1年遡るごとに-1となるように指定した。

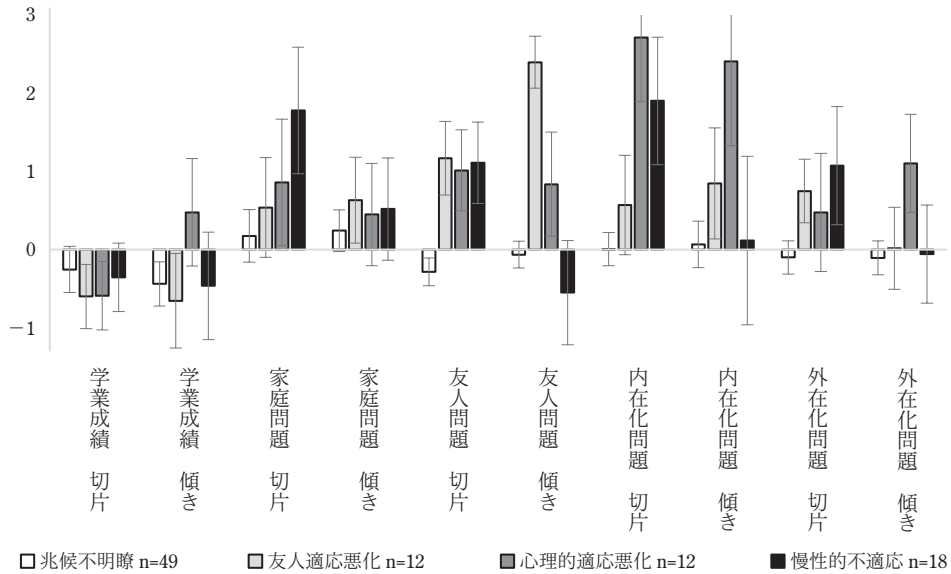
成長混合分布モデルでは，これらの成長因子（切片と傾き）の分布が，未知の下位集団の分布の組み合わせによって形成されていると仮定する。図には学業成績のみを示したが，これに加え不登校の発生を予測すると考えられる家庭問題，友人関係問題，内在化問題，外在化問題を同時に分析に使用した。学業成績については学校で実施されている学力テストの記録，その他の変数は質問紙尺度の得点（児童生徒，保護者，教師の評定の合成得点）を分析に用いた。これら5つの変数の小4から中1にかけての変化の軌跡を，成長混合分布モデルにより複数の類型に分けることを試みた。ただし，一度に全ての変数を分析に使用すると推定が著しく不安定になったた

め，初めに個々の変数ごとに成長曲線モデルにより切片と傾きの因子得点を得て，それらを用いて混合分布モデルによる推定を行った。分析には，中2時点で初めて不登校が見られた生徒（5つの生年コホートの91名）と不登校を経験していない生徒（3325名）のデータを使用した（不登校の新たな発生を予測する目的のため，中1までの段階ですでに不登校を経験している生徒のデータは除外した）。

不登校生徒91名の心理社会的適応の軌跡について，成長混合分布モデルによる類型化を行った結果を Figure 3 および Figure 4 に示す。適合度指標，分類確率（各個人の各クラスに対する割り当ての正確さ），人数比（極端に人数の少ないクラスがないか），解釈可能性などを総合的に考慮し，4クラス解を採用した。「兆候不明瞭群」と命名したグループ（ $n=49$ ）は，学業成績の傾きが全体平均を有意に下回することを除いては，不適応の兆候が見られなかった。このグループは不登校生徒の半数以上（54%）を占めた。「友人適応悪化群」（ $n=12$ ）は，友人問題の傾きが突出して高いことが特徴的であり（Figure 3），Figure 4 から小4から中1にかけて，友人関係適応が顕著に悪化してきていることがわかる。「心理的適応悪化群」（ $n=12$ ）は，心理的不適応を表す内在化問題の切片・傾きや外在化問題の傾きが高く（Figure 3），心理的適応が悪化する傾向にあることが読み取れる（Figure 4）。「慢性的不適応群」（ $n=18$ ）は，家庭問題，友人問題，内在化問題，外在化問題の切片が高い値を示す一方，各変数の傾きはいずれも平均的水準に留まっており（Figure 3），小4から中1にかけて，幅広い領域で慢性的に不適応の状態にあることがうかがわれる（Figure 4）。

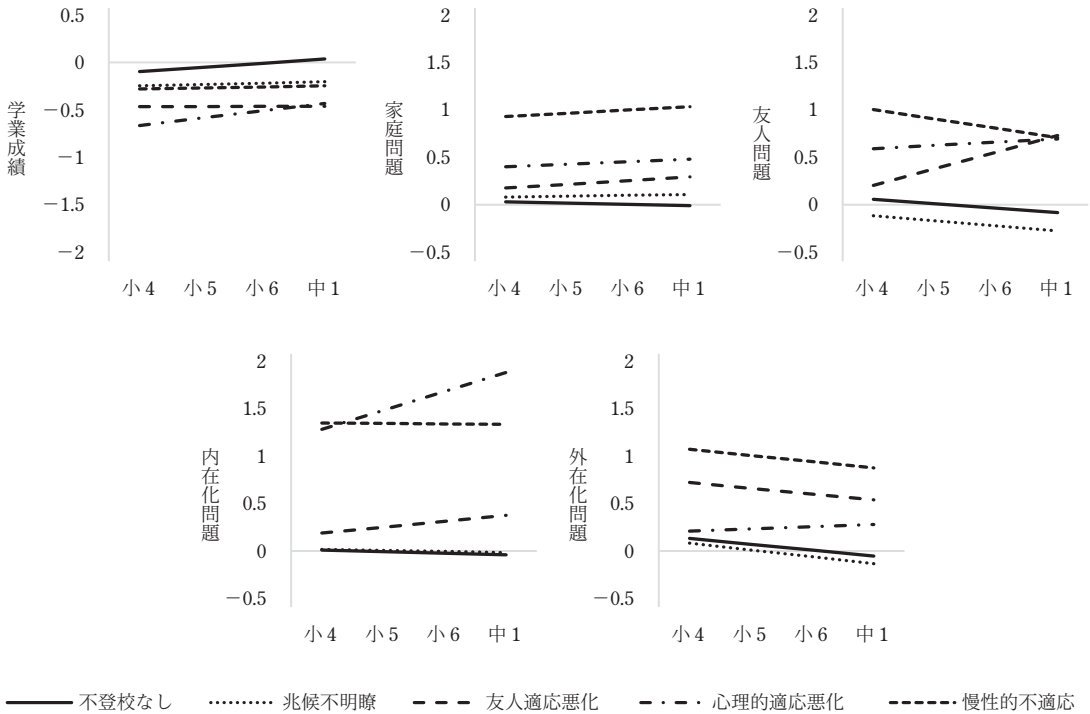
類型ごとの不登校の発生の予測

これらの結果から，不登校生徒の心理社会的適応の軌跡には複数の異なるパターンが存在することが示された。こうしたパターンのうち，明確な不適応の兆候が見られない「兆候不明瞭群」については前年度までの心理社会的適応の軌跡によって不登校の発生を予測することは難しいかもしれないが，固有の特徴を示した他の3群については前年度までの段階で，ある程度の予測を行うことができると考えられる。そこで，4つの類型ごとに，心理社会的適応の軌跡によって不登校の発生をどの程度予測しうるかを多項ロジスティック回帰分析によって検証する。多項ロジスティック回帰分析は，複数のカテゴリを持つ質的変数を従属変数とした回帰分析であり，特定のカテゴリ（ここでは不登校を経験していない生徒）を基準としたときの，他の各カテゴリ（ここでは不登校の4類型）に対するオッズ（不登校のように生起頻度が低い事象の場合，確率の近似値と見なせる）が，各説明変数の値によってどのように変化するかを推定す



注. グラフは各成長因子の標準化得点の平均値, エラーバーは95%信頼区間を示す。エラーバーが0をまたがなければ全体平均と有意差があることを意味する。

Figure 3 成長混合分布モデルによる不登校生徒の軌跡の類型化



注. 解釈を容易にするため各変数はあらかじめ標準化。

Figure 4 各類型における心理社会的適応の変化の軌跡の推定値

Table 1 不登校の4類型に対する多項ロジスティック回帰分析の結果

	兆候不明瞭群				友人適応悪化群				心理的適応悪化群				慢性的不適応群			
	モデル1		モデル2		モデル1		モデル2		モデル1		モデル2		モデル1		モデル2	
	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p
学業成績 切片	0.71	.022	0.61	.001	0.69	.160	0.45	.009	0.70	.267	0.88	.768	0.99	.953	0.77	.317
家庭問題 切片	1.34	.057	1.20	.283	1.17	.589	0.97	.939	1.10	.718	1.12	.732	2.25	.000	2.08	.002
友人問題 切片	0.52	.007	0.56	.025	1.82	.023	0.67	.422	0.47	.021	0.35	.024	0.60	.049	0.77	.343
内在化問題 切片	1.31	.167	1.30	.194	0.84	.583	0.63	.286	5.87	.000	6.78	.000	2.77	.000	2.68	.000
外在化問題 切片	0.86	.404	0.84	.351	1.16	.561	1.09	.790	0.89	.719	0.73	.499	1.47	.069	1.46	.091
学業成績 傾き			0.58	.000			0.41	.019			1.78	.171			0.50	.013
家庭問題 傾き			1.24	.190			1.59	.196			0.89	.725			0.95	.839
友人問題 傾き			1.03	.883			9.09	.000			1.17	.600			0.67	.040
内在化問題 傾き			1.02	.886			1.38	.356			1.57	.095			0.88	.516
外在化問題 傾き			0.83	.237			0.62	.112			3.26	.001			0.93	.779
Nagelkerke R ²	.035		.068		.087		.455		.368		.518		.299		.358	

注. ORは各独立変数が1上昇したときのオッズ比を表す。不登校なし群を基準カテゴリーとした。解釈を容易にするため各独立変数はあらかじめ標準化した。モデル1は各要因の切片のみ、モデル2は切片と傾きを独立変数としたモデル。

ることができる。つまり、ここでは不登校の4類型ごとに、心理社会的適応の軌跡によって不登校の発生リスクをどの程度予測しうるかを推定することになる。また、切片因子（中1時点の水準）と傾き因子（小4から中1にかけての変化）のそれぞれがどの程度の説明力を持つかを検証するために、まずモデル1として切片因子のみを投入し、その後、モデル2として傾き因子を含めた場合の説明力の変化を確認する。

その結果をTable 1に示した。「兆候不明瞭群」では、学業成績の切片や傾きが1を有意に下回るオッズ比を示している。例えば学業成績の切片は0.61のオッズ比を示していることから、学業成績の切片が1SD上昇するごとに、「兆候不明瞭」型の不登校に陥るリスクが0.61倍に減少することが示された。同様に、学業成績の傾きもこのタイプの不登校のリスクを押し下げることが示された。一方、友人問題の切片が同様に負の効果を示したことは想定外の結果である。「兆候不明瞭群」は、前年度までには明確な不適応が生じていないが、不登校の発生年度にかけて何らかの不適応状態が発生することで不登校につながるパターンとして理解できる（実際に当該年度の適応を分析すると、それを裏づける結果が得られる）。そう考えれば、前年度に友人関係適応が良好であるほど、そこから不適応状態に陥ったときの落差が大きくなり、不登校のリスクを高める可能性が考えられる。モデルの説明力は、切片因子のみの場合が.035、傾き因子を含めた場合が.068で、傾き因子により説明力が2倍程度に上昇しているが、いずれも絶対的な水準としては小さな値であり、このタイプの不登校の発生を前年度までの適応の軌跡により高い精度で予測することは困難であると言える。

「友人適応悪化群」については、学業成績の切片・傾きが負の効果を示すとともに、友人問題の傾きが9.09という顕著なオッズ比を示している。友人関係適応の悪化は、ある種の不登校を予測する上で特に重要性の高い因子であることが示された。モデルの説明力は、傾き因子を含めることで.087から.455まで飛躍的に上昇しており、一時点（前年度）の水準だけでなく、過去数年の変化量に着目することの重要性が明らかとなった。

「心理的適応悪化群」については、友人問題の切片が負の効果を示す一方、内在化問題の切片や外在化問題の傾きが正の効果を示し、内在化問題の傾きも有意傾向の正の効果を示した。上記の友人関係の悪化と同様、心理的適応の悪化もまた、不登校を予測する重要な因子となることが示された。モデルの説明力は傾き因子により.368から.518まで上昇しており、ここでも変化を考慮に入れることの重要性が示唆された。

「慢性的不適応群」に関しては、家庭問題、内在化問題の切片が有意な正の効果を示すとともに、学業成績、

友人問題の傾きが負の効果を示した。モデルの説明力は、傾き因子を含めても.299から.358に上昇するに留まり、このタイプの不登校の予測には傾き（変化）よりも切片（水準）が重要な役割を果たすことが示唆された。

以上の検証により、主に以下の3点の知見が得られた。①不登校が発生するプロセスには、友人関係適応の悪化、心理的適応の悪化、慢性的な不適応、直近の突発的な不適応など、質的に異なる複数のパターンがある。②こうしたパターンの違いを考慮すれば、（中2時点で生じる）不登校の半数程度は、前年度までの心理社会的適応の軌跡によって比較的高い精度で予測できる。③前年度の状態だけでなく、過去数年の変化に着目することで、一部の不登校の予測精度を大幅に高められる。こうした不登校の予測システムを学校現場に実装するには、変数・人数の両面でより豊富なデータを得て類型の精緻化を図るとともに、交差妥当化により結果の再現性を検証することが必要となるが、児童生徒の適応状態の縦断的な軌跡のモニタリングが不登校の発生リスクの推定に貢献しうることを示唆する今回の知見は、不登校を始めとする深刻な情緒・行動問題の予防方策を検討していく上で、一つの有望な方向性を示すものと言えようである。

このように変数の経時的な個人内変動の軌跡とその個人差や関連を検証できることは、縦断研究に特有の強みである。ここでは変化の軌跡の個人差を表現する方法や、そうした軌跡の個人差によって特定の事象の発生を予測する方法について論じたが、発達研究においては、ある事象や特性によって発達の軌跡がどのように変化するかという視点を持つことも重要である。例えば、幼少期に逆境を経験した個人の発達軌跡が個人要因や環境要因によってどのように調整されるかというレジリエンス研究や、心理学的介入によって長期的な適応の軌跡がどのように変化するかという介入研究などである。

因果関係の推論

心理学領域において縦断研究が用いられるとき、その目的の最も大きな割合を占めるのが因果関係の推論であると思われる。一般に、心理学研究の目的は、(1)記述、(2)予測、(3)説明、(4)制御という4つに大別されるが(Coon & Mitterer, 2013)、実際には(3)、つまり変数間の因果関係の探索や検証に主眼を置いた研究が大半を占めているためであろう。因果関係に関する手がかりが得られることは、前項で論じた個人内の変化の軌跡を追うことができるという縦断研究の強みから副次的に生じるメリットであるが、一部では、これが縦断研究の唯一のメリットであると誤解されているようにも見受けられる。本稿の冒頭に述べたように、横断研究の論文

において、しばしば今後の展望として縦断研究の必要性が述べられるのも、(明示的には記載されていないことも多いが) 因果関係について直接的な検証ができるという認識に基づいているものと考えられる。

確かに縦断研究は実験的操作をとみなわない観察研究において、因果関係に関する手がかりを得るための有効な手段の一つではあるが、唯一の手段ではない。近年、観察データに基づいて統計的に因果関係の推論を行うための準実験 (quasi-experiment) の方法論が飛躍的に発展を遂げており、代表的な方法として、傾向スコア法 (Rosenbaum & Rubin, 1983)、操作変数法 (Angrist, Imbens, & Rubin, 1996)、回帰分断デザイン (Thistlethwaite & Campbell, 1960) などがある。縦断研究は、横断研究と比較すれば、因果関係に一步近づくことのできる方法であるが、こうした緻密な準実験デザインと比較したときには、その効力は限定的である場合も多いことを認識しておく必要がある。

因果関係を証明するためにどのような要件が必要となるかは科学哲学において古くから議論されてきたが、少なくとも John Stuart Mill が論じた以下の3点を示す必要があるという考え方が一般的である (Cook, Campbell, & Shadish, 2002)。

- (1) 仮説上の原因と結果が相関している
- (2) 仮説上の原因が結果よりも先に生じている
- (3) 観測された相関に関する他の説明 (i.e. 交絡因子による疑似相関である可能性) を排除できる

このうち、(1) については横断研究によっても確認することができる。(2) については、一時点の横断研究では示すことが難しいが、複数時点でデータを得る縦断研究であれば、少なくとも限定的な証拠を得ることが可能である。一方、(3) については、観察研究である限り、横断研究、縦断研究のいずれにおいても、(限りなく蓋然性を高めていくことはできるが) 決定的な証拠を示すことは難しい。つまり、因果関係の推論に関して縦断研究が横断研究に明確に優っているのは、基本的に (2) の一点においてのみである (3) に対してもメリットを提供するという議論はあるが、それについては後述す

る)。以下、縦断研究において (2) および (3) を示す上で方法論的な問題について詳細に論じていく。

時間的順序性

交差遅延相関と交差遅延回帰 上述のように、因果推論における縦断研究の主な貢献は、仮説上の原因と結果の時間的順序性について手がかりを得ることができるという点にある。横断研究においては、Figure 5の左図のように2つの変数が測定されても、得られる相関係数は1つであり、友人問題→内在化問題の効果と内在化問題→友人問題の効果を区別して推定することはできない。なお、横断データに対してパス解析などの因果関係を扱う分析手法が適用されることがあるが、この場合、モデル上の因果関係の方向性は何らかの理論的あるいは実証的根拠により「仮定」されるものであって、データにより「検証」されるものではないことに注意が必要である (この点に関する誤解も非常に多く見られる)。

これに対し、縦断研究では各変数について複数時点でデータを得ることができるため、Figure 5中央の図のように、1時点目の各変数の値が2時点目の他方の変数の値を予測するか否かを検証することができる。この交差遅延相関モデルは、一見すると、各変数の測定時点間に時間差が存在するため、時間的順序性を担保した上で双方向の効果も推定することができるように思われる。しかし、実際には、この方法では時間的順序性を十分に確保することができていない。なぜなら、2変数の間にはTime 1の時点ですでに相関があり、その相関が一定程度Time 2に引き継がれていると考えられるためである。例えば、Time 1の友人問題とTime 2の内在化問題の相関には、直接の因果的效果とは別に、Time 1の内在化問題を媒介した相関が含まれる。

この問題を解消するためには、Figure 5右図のように、Time 1時点での変数間の相関と2時点間での同一変数間の効果を考慮し、統計的に調整する必要がある。これにより、例えば、Time 1の友人問題からTime 2の内在化問題の効果は、Time 1時点の内在化問題を一定にしたときの効果を表すことになる。言い換えれば、Time 1の友人問題の水準が、Time 1からTime 2にか

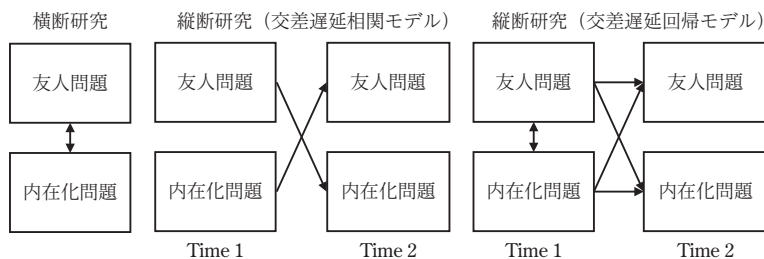


Figure 5 因果推論のための分析モデルの模式図

けての内在化問題の変化を予測するかを検証していることになる。このとき、Time 1からTime 2にかけての内在化問題の変化は、Time 1における友人問題の測定よりも必然的に後に生じているため、両者の間には時間的な順序性があると見なすことができる。これが発達心理学領域の縦断研究において最も広く用いられる交差遅延回帰モデル（Cross-Lagged Panel Model: CLPM）である。前項の不登校の分析において、中1までの段階ですでに不登校を経験している生徒を分析から除外したのも、中1から中2にかけての不登校の新たな発生のみを扱うことで、中1までに測定された説明変数との時間的順序性の明確化を図ったものであった。このように、目的変数の従前の状態を統制するというアプローチは、縦断研究において説明変数と目的変数の時間的順序性を確保するための基本的戦略である。

交差遅延回帰モデルの拡張 しかし、近年、CLPMでも時間的順序性を十分明確にはできていないとして、いくつかの新たなモデルが提唱されている（Orth, Clark, Donnellan, & Robins, 2021）。その中でもHamaker, Kuiper, & Grasman (2015)のランダム切片交差遅延モデル（RI-CLPM; Figure 6）は発達領域でも広く普及しつつある。Hamaker et al. (2015)は、マルチレベル分析の発想を援用し、縦断データにおける各観測変数の値には、各個人の安定的な特性（個人間変動）と時点ごとの変動（個人内変動）の両方が含まれており、両者を明確に分離した上で変数間の関連を検証する必要があると主張している。こうした考えに基づき、Figure 6のように、個人間変動を表す特性因子（ランダム切片）を導入し、それによって説明されない個人内変動（残差）に

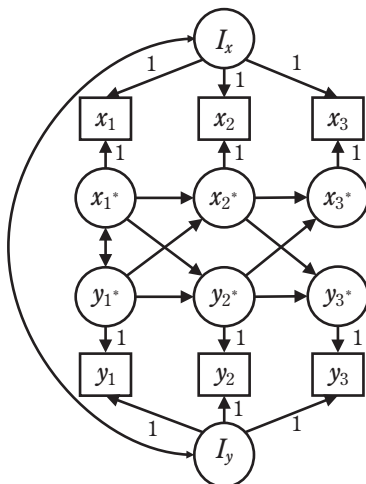


Figure 6 RI-CLPMの模式図

おいて交差遅延効果を設定している。Hamaker et al. (2015)によれば、従来のCLPMでは、個人間変動と個人内変動が分離されずに交差遅延効果が推定されているため、この推定値には個人間レベルでの関連と個人内レベルでの関連が混在しているとされる。これを裏付けるため、Lucas (2022)はシミュレーションによる検証を行っている。個人間の関連のみを想定したFigure 7左のモデルにより生成したデータに対してCLPMを適用した結果、右図のように疑似的な交差遅延効果が見出された。このように、安定的な個人差が存在する変数について通常のCLPMを適用した場合、推定された交差遅延効果には、時間的順序性の不明確な個人間レベルの関連が含まれる可能性がある。

一方で、CLPMとRI-CLPMは互いに異なる側面に焦点を当てたモデルであり、単純な優劣の関係にはないため、研究目的によって使い分ける必要があるという指摘もある（Lüdtke & Robitzsch, 2021; Orth et al., 2021）。Lüdtke & Robitzsch (2021)は、CLPMが一方の変数が1単位上昇したときの他方の変数への効果を推定するのに対し、RI-CLPMは個人平均からの偏差が1単位上昇したときの効果を推定していることを指摘し、後者は一時的な変動のみを捉えているため、典型的な縦断研究の目的にはそぐわないと述べている。

縦断研究のリサーチクエストは多様であり、例えば日誌法を用いた研究で、ある日の出来事が翌日の気分によどのような影響を及ぼすかといった問いであれば、RI-CLPMにより安定的な個人差を取り除いた上で、個人内変動のみを扱うことが望ましいと考えられる。fMRIや脳波などの生理学的指標の時系列的解析においても、個人間変動と個人内変動を明確に分離した上で分析を行うことが一般的である。こうした短期間のインテンシブな縦断データに対してはRI-CLPMの発想はなじみやすい。一方、数カ月から数年の間隔で行われる縦断

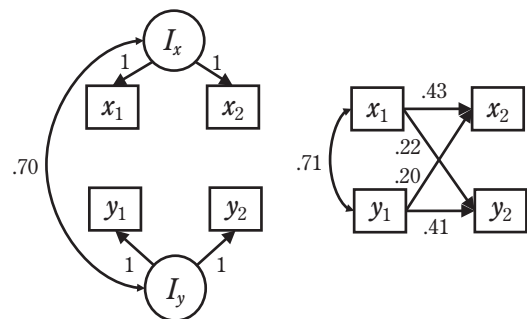


Figure 7 データ生成モデル（左）とCLPMの推定結果（右）（Lucas, 2022より一部改変）

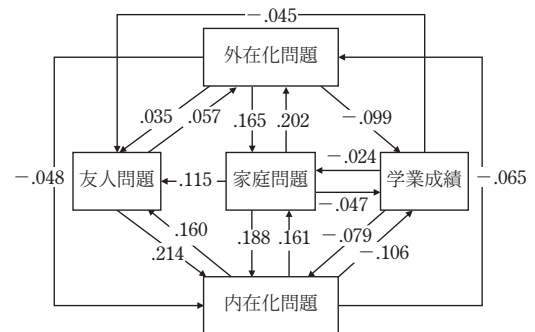
研究では、Lüdtke & Robitzsch (2021) も指摘するように、個人平均の周囲で揺れ動くような一過性的変動よりも、漸進的・累積的な変化のプロセスに主眼が置かれており、特性を取り除いた残差間の関連を扱うという RI-CLPM の発想が適さないケースが多いと思われる。また、長い調査期間においては個人の特性そのものが時間的に変化していくことが想定されるため、調査期間内の平均を各個人の安定した水準と見なして機械的に個人内変動と切り分けることにも無理があるように思われる。近年の遺伝学においても、エピジェネティクスの概念に代表されるように、遺伝的要因による個人差は、初めから安定的に存在するものではなく、環境との相互作用の中で動的に変化するものと考えられており、RI-CLPM が想定する特性因子という静的な個人差の概念には適合しない可能性がある。

しかし、Hamaker et al. (2015) や Lucas (2022) が指摘したように、個人レベルの関連が交差遅延効果に混入するという CLPM の問題は、時間的順序性という因果推論の基本的前提を脅かすものであり重要である。この点に関して、Lüdtke & Robitzsch (2021) は、1時点前からの効果のみを想定する CLPM に対して、2時点前からの遅延効果を追加することによって、個人レベルの関連によるバイアスある程度抑制できることを示している。原理的には、2時点前に限らず、より高次の遅延効果を含めることで、バイアスをさらに低減することができると考えられる。また、Lüdtke & Robitzsch (2021) は、2次の遅延効果を含む CLPM (CL2PM) をデータ生成モデルとした場合、RI-CLPM が過剰調整によるバイアスを生じさせることを示した。これは RI-CLPM を生成モデルとした場合に、CLPM が不正確な結果をもたらすことを示した前述の Lucas (2022) とは対照的な知見である。さらに Lüdtke & Robitzsch (2021) は、CL2PM と RI-CLPM のいずれが妥当であるかをデータへの適合度に基づいて区別することはできないことを指摘している。こうした知見を踏まえれば、データの生成メカニズムが未知である場合、特性因子が全ての時点の観測変数に等しく影響することや特性因子と個人内変動が無相関であることなどの強い仮定を置く RI-CLPM よりも、比較的シンプルな仮定のもとで個人内変動をある程度調整できる高次の遅延効果を設定した CLPM を用いた方が、バイアスの小さい推定値を得やすいという考え方もありうる。異なる見方として、本特集の宇佐美 (2022) の寄稿も参照されたい。

成長曲線モデルの応用 もう一つの方法として、時間的順序性の担保が難しい個々の時点間の関連を扱う CLPM や RI-CLPM ではなく、前項で紹介した成長曲線モデルを用いて、一貫した変化の傾向を扱うというアプローチもある。例えば、友人問題と内在化問題のそれぞ

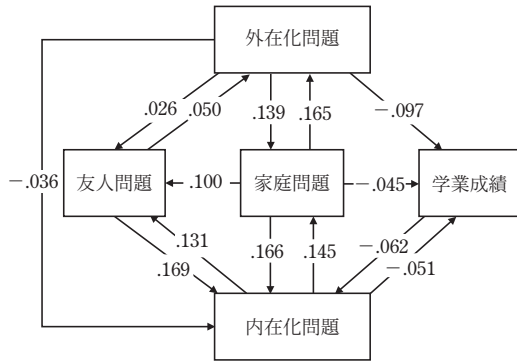
れについて、小4から中3の6時点のデータを用いて切片(小4時点の水準)と傾き(小4から中3の変化の傾向)を推定し、各変数の切片が他方の変数の傾きをどう予測するか(当該変数の切片を調整した上で)を検証するといった方法である。この方法では、時点間の一時的変動の情報を捨象する代わりに、調査期間全体を通しての一貫した変化の傾向を扱うため、CLPM で生じたような個人間変動と個人内変動の交絡の問題を避け、時間的順序性をより明確化することができる。一方で、結果の持つ意味合いが CLPM とはやや異なる点に注意が必要である。CLPM の交差遅延効果は、ある時点での説明変数が1単位上昇したときに、その時点から次の時点にかけて目的変数がどの程度上昇するかを表しているが、成長曲線モデルにおける切片→傾きの効果は、説明変数の初期値が1単位上昇したときに、調査期間全体を通じた目的変数の変化傾向(の調査間隔あたりの速度)がどの程度上昇するかを表す。

心理社会的適応の要素間の縦断的関連 著者らの研究データを用いて、心理社会的適応に関する5変数(学業成績、家庭問題、友人問題、内在化問題、外在化問題)の間の縦断的関連について、小4から中3の6時点において CLPM, CL2PM, 成長曲線モデルによる推定を行った結果を Figure 8~Figure 10 に示す。比較のために横断データ(中3時点)における5変数間の相関係数も Table 2 に示した。分析には全時点で測定値が得られている4コホートの3,018名のデータを使用した。CLPM, CL2PM については、各変数の安定性(同一変数の時点間の自己回帰係数)が高いほど交差遅延効果が低く推定されやすい性質があるため(Adachi & Willoughby, 2015)、相互の比較を容易にするため、小4から中3にかけての効果の累積値を推定した。例えば、学業成績→家庭問題の効果については、以下の5つの効果



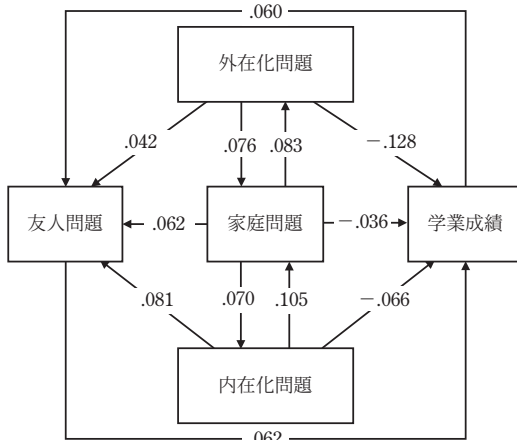
注. 小4から中3にかけての効果の累積値。あらかじめ全ての変数を標準化。有意な効果のみ図示。

Figure 8 CLPM による交差遅延効果の推定結果



注. 小4から中3にかけての効果の累積値。あらかじめ全ての変数を標準化。有意な効果のみ図示。

Figure 9 CL2PMによる交差遅延効果の推定結果



注. あらかじめ全ての変数を標準化。有意な効果のみ図示。

Figure 10 成長曲線モデルによる切片→傾きの推定結果

(4つの間接効果と1つの直接効果)の合計値を推定した (CL2PMでは2次の遅延効果を介した効果も含めた)。

小4 学業→小5 家庭→小6 家庭→中1 家庭→中2 家庭→中3 家庭

小5 学業→小6 家庭→中1 家庭→中2 家庭→中3 家庭

小6 学業→中1 家庭→中2 家庭→中3 家庭

中1 学業→中2 家庭→中3 家庭

中2 学業→中3 家庭

この値は、小4から中2にかけて説明変数の値が継続して1SD高かったとき、中3時点の目的変数の値が何SD高くなる(または低くなる)かを表す。一方、成長曲線モデルによる切片→傾きの推定値は、小4時点での説明変数の初期値が1SD高かったとき、小4から中3にかけての目的変数の変化の傾きが何SD高くなるかを意味する。なお、分析に際しては、性別およびコホートを統制変数として調整した。

CLPMとCL2PMの結果を比較すると、全体として顕著な結果の違いはないが、2次の遅延効果が調整されたCL2PMの推定値がいずれも小さくなっており、CLPMで見られたいくつかの効果がCL2PMでは有意でなかった。全体として横断データの相関とCLPMやCL2PMにおける交差遅延効果は類似した傾向を示しているが、顕著な違いとして、内在化問題と外在化問題は、横断データでは.377という正の相関を示しているが、交差遅延効果では弱い負の効果に転じている。友人問題と外在化問題も、横断データでは.368という比較的明確な正の相関を示しているが、交差遅延効果では.05前後の弱い効果に留まっている。CLPMやCL2PMでは、双方向の交差遅延効果を示す変数対が多かったが、家庭問題→友人問題と外在化問題→学業成績では、単方向の効果が見られた。

成長曲線モデルによる推定においても、CLPMやCL2PMとおおむね類似した結果のパターンが見られたが、明確な違いとして、双方向の効果を示す変数対が3組に留まり、他の6組の変数対では単方向の効果のみが

Table 2 横断データ (中3時点) における5変数間の相関係数

	学業成績	家庭問題	友人問題	内在化問題	外在化問題
学業成績					
家庭問題	-.129				
友人問題	-.150	.239			
内在化問題	-.181	.322	.600		
外在化問題	-.226	.323	.368	.377	

注. 係数は全て0.1%水準で有意。

見られた。一つの可能性として、Lucas (2022) のシミュレーションが示したように、CLPM や CL2PM では個人間レベルの関連が疑似的に双方向の交差遅延効果として検出されていた可能性が考えられる。もう一つの可能性としては、実際に CLPM や CL2PM が扱う個々の時点間の関連では双方向の効果が生じていたのに対し、成長曲線モデルが扱う初期値と調査期間全体の変化傾向の関連では単方向の効果しか生じていなかったということも考えられる。例えば、友人問題→内在化問題の効果は、成長曲線モデルでは見られなかったが、CLPM や CL2PM ではモデル中で最も高い推定値を示している。友人関係上の問題が子どもの主要なストレスの一種であることについては、多くの理論的・実証的根拠が蓄積されてきており (Platt, Kadosh, & Lau, 2013)、個々の時点間の短期的な関連としてこの効果を見出した CLPM や CL2PM の結果には十分な妥当性があると考えられる。一方で、成長曲線モデルでこの効果が見出されなかったのは、このモデルにおける切片 (初期値) が、友人関係に関する個人の安定的な特性 (いわば内向性) を色濃く反映しており、その特性の水準自体は内在化問題を高める方向には働かないということを示唆していると考えられる。そのように考えると、成長曲線モデルにおいて見られた (一見不可解な) 友人問題→学業成績の相互の正の効果も、内向的な特性が (おそらく学習時間の相対的な多さを媒介して) 学業成績の向上につながることを反映していると考えられる。

以上のように、個々の時点間の短期的関連を扱う CLPM や CL2PM と調査期間全体の長期的関連を扱う成長曲線モデルでは、同一のデータを用いた分析でも異なる結果をもたらす可能性があり、その結果が持つ意味合いも質的に異なる。研究目的に応じて両者を使い分けるか、上のように両者を組み合わせて多層的な理解を得ることも有用である。

交絡因子の統制

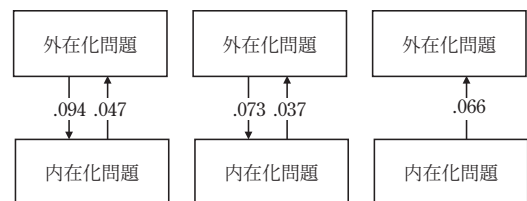
因果関係を証明する上でのもう一つの要件は、観測された相関が交絡因子による疑似的な相関でないことを示すことである。交絡因子とは、仮説上の説明変数と目的変数の双方に影響する共通要因である。交絡因子の説明変数に対する効果を a 、目的変数に対する直接効果 (説明変数を統制したときの効果) を b としたとき、交絡因子が統制されなければ、説明変数の目的変数に対する効果の推定値には $a \times b$ の疑似相関が含まれることになる。

因果関係の検証において実験研究が最も強力なエビデンスを提供すると考えられているのは、時間的順序性の要件に加えて、交絡因子の統制においても特別なメリットを提供するためである。時間的順序性に関して言えば、実験研究では、要因の実験的操作が行われた後にど

のような変化が生じるかを測定するため、仮説上の原因と結果の時間的順序性が明確である。交絡因子の統制という点では、各参加者への条件の割り当てを無作為に行うランダム割付によって、各条件の参加者の性質に系統的な偏りが生じる可能性を排除している。それでも実験手続き上の問題による交絡が生じる可能性は除外できないが、少なくとも参加者レベルの交絡因子は (偶発的な偏りを除いて) ほぼ完全に統制することができる。この方法の強みは、どのような交絡因子が存在するかを具体的に検討することなく、参加者レベルの全ての交絡因子の影響を一挙に遮断することができる点にある。

一方、実験操作をとまなわない観察研究では、こうした強力なメリットを持つランダム割付を用いることができないため、何らかの代替手段によって交絡因子の統制を図る必要がある。前述の準実験は、その中でも比較的頑健性が高いと考えられている手法であり、観察データにおいても実験と類似した状況を作り出すことによって因果関係を示そうとするものである。それに対し、一般的な縦断研究において用いられる方法は、横断研究と同じく、回帰モデルによる交絡因子の統計的調整である。この方法は、交絡因子となるであろう変数を説明変数や目的変数とともに測定しておいて、重回帰分析や構造方程式モデリングにおける共変量としてモデルに含めるというものである。

交絡が推定結果にどのような影響を与えるかを示すため、前述の CLPM, CL2PM, 成長曲線モデルによる分析について、内在化問題と外在化問題の2変数のみを使用した場合の推定結果を Figure 11 に示す。CLPM, CL2PM で双方向の正の効果、成長曲線モデルで単方向の正の効果が見られた。一方、5変数を用いた分析では、CLPM (Figure 8) や CL2PM (Figure 9) で内在化問題と外在化問題の間に双方向または単方向の負の効果が見られ、成長曲線モデル (Figure 10) では有意な効果が見られなかった。



注. CLPM と CL2PM の結果は小4から中3にかけての効果の累積値。あらかじめ全ての変数を標準化。有意な効果のみ図示。

Figure 11 2変数のみを用いた場合の CLPM (左), CL2PM (中央), 成長曲線モデル (右) による推定結果

こうした結果を比較すると、2変数間の分析でのみ見られた相互の正の効果は、他の3変数との交絡による疑似的な関連を反映していると推察できる。内在化問題と外在化問題はいずれも心理的適応に関する変数であるのに対し、残りの3変数（家庭問題、友人問題、学業成績）は社会的適応に関する変数である。このうち家庭問題と友人問題は、Figure 8~Figure 10より、内在化問題と外在化問題の双方に正の効果をもたらす交絡因子となっていることがわかる。これにより、家庭問題や友人問題を統計的に調整しなかった場合には、内在化問題と外在化問題の間に正の疑似的な関連を生じさせたものと考えられる。

このように、縦断データを用いた分析でも横断データを用いた分析と同様に、交絡因子による疑似相関の問題は生じうる。そのため、本稿の冒頭に述べたように、2変数のみを経時的に測定してCLPMなどによる分析を行ったとしても、交絡因子が存在しないことの根拠が明示されていなければ、因果関係を示すエビデンスとしては限定的な効力しか持たないことに注意が必要である。現実的には、交絡因子が存在しないというケースは想定しにくいいため、関心対象の目的変数に影響を及ぼすと考えられる変数は、（その効果自体に関心がなくても）可能な限り幅広く測定しておいて、統計的に調整を図ることが望ましい。また、性別、年齢、地域などの人口統計学的変数も重要な交絡因子となることが多いため、一通り共変量として投入する必要がある。

こうした認識は、経済学や疫学の領域では広く共有されており、事前の検討に基づいて特定された多数の共変量を用いて分析が行われることが一般的である。一方、心理学領域の論文では関心対象の変数のみが扱われ、交絡因子の理論的検討はもとより、基本的な人口統計学的変数の調整さえ行われていないケースも多い。パス解析や構造方程式モデリングの手法が普及したことで、媒介変数や調整変数を含む複雑な仮説モデルを想定した分析が多く行われるようになった一方で、モデル中で個々の「パス」として表現される各変数間の因果関係を検証することへの意識が薄れているようにも見受けられる。ランダム割付をとまなう実験デザイン（もしくは、それに準じた方法で交絡因子の統制を図る準実験デザイン）を用いない限り、交絡因子による疑似相関を防ぐには、研究開始前の段階で理論的・実証的根拠に基づいて交絡因子を同定した上で、それらを個別に測定し、統計的に調整を図る必要があり、それは縦断研究であっても例外ではない¹⁾。こうした基本的な認識を心理学領域においても広く共有する必要があるように思われる。

ただし、交絡因子を時間変動的 (time-variant) な因子と時間不変的 (time-invariant) な因子に分けたとき、RI-CLPMのように個人内変動と個人間変動を分離する

アプローチを用いれば、時間不変的な因子については統制が可能であるという見方もある (Bailey, Oh, Farkas, Morgan, & Hillemeier, 2020)。また、経済学分野におけるパネルデータ解析では、RI-CLPMと同様に個人ごとの平均的水準を考慮したモデルが広く用いられており、その利点は時間不変的な交絡因子を統制できることにあるとされている (Wooldridge, 2010)。しかし、前節でも述べたように長期間の変動と関連を扱う発達研究において、個人内平均 (特性因子) の導入という単純なアプローチで個人内変動と個人間変動を明確に分離しうることについては議論の余地がある。また、RI-CLPMを用いたシミュレーションでも、観測されていない時間不変的な交絡因子の統制には限界があることが示唆されている (Lüdtke & Robitzsch, 2021)。こうした議論を踏まえると、交絡因子の統制については、RI-CLPMなどの統計モデルに依存するよりも、事前に十分な理論的・実証的根拠をもとに交絡因子を絞り込み、それらを個別に観測して調整を図るという古典的な方法を用いた方が、泥臭いながらも堅実であるように思われる。

まとめ

本稿では、発達心理学における研究の文脈に即して、縦断研究の縦断研究の What, Why, How について議論した。What に関して言えば、縦断研究とは同一の対象から複数時点のデータを収集し、個人内の時系列的な変化や関連に着目した分析を行う研究の総称であり、数か月から数年の間隔で追跡的に調査を行う典型的な縦断研究だけでなく、後方視的研究、生理学的研究、経験サンプリング研究、単一事例実験、長期縦断研究など、きわめて多様な研究デザインが含まれる。研究対象の構成概念やリサーチクエスションの性質に応じて、適切な研究デザインを採用する必要がある。

Why に関して言えば、縦断研究の意義は、(1) 年齢・コホート・時期の効果を分離できる、(2) 個人内変動の軌跡とその個人差を定量化できる、(3) 因果関係に関する手がかりを得られるという3点に集約される。このうち、(1) や (3) は他の研究デザインによっても代替が可能であるが、(2) は同一個人内のデータの紐づけをとまなう縦断研究に固有の強みである。つまり縦断研究は、変化の軌跡にどのような個人差があるか、変化の軌跡を何が予測するか、変化の軌跡によって何が予測されるか、複数の変数の変化の軌跡に関連があるか、といった発達心理学上の重要な問いに答えを与えることのできる唯一のアプローチである。

こうした縦断研究の視点の複雑さは、研究デザインや

1) 準実験においても傾向スコアを用いた方法では交絡因子の測定が必要となる。

データ分析などのHowの複雑さに直結する。上記の(1)の検証のためには、データの階層性の問題や識別不定の問題に対処する必要があるが、現時点でアルゴリズム的な解決策は存在せず、個々のデータの性質に応じて試行錯誤的な対応が必要となる。(2)の検証には、本稿で用いた潜在成長曲線モデルなどの構造方程式モデリングに基づく方法の他に、マルチレベルモデルに基づく分析手法も存在し、研究目的に応じて適切に使い分ける必要がある。混合分布モデルも発展途上の分析手法であり、推定法やクラス数の決定など手探りで進めなければならない部分が多い。こうした分析手法を適用できるような研究のデザイン(説明変数、目的変数、時点数、測定タイミングなど)をあらかじめ設定しておくことも不可欠となる。(3)の検証は最も困難な課題である。要件の一つである要因間の時間的順序性を検証するためには、CLPM、CL2PM、RI-CLPM、成長曲線モデルなどの分析手法の性質を理解した上で、研究の文脈に適した方法を採用する必要がある。もう一つの要件である交絡因子の統制のためには、横断研究と同様に、あらかじめ候補となる交絡因子をできる限り網羅的に同定した上で調査計画に組み込んでおくことが必要となる。また、因果推論における縦断研究の限界を認識し、実験法や準実験など、因果関係の検証により適したデザインを採用することも選択肢の一つである。

縦断的アプローチにより、個人内の時間的変動という軸が加わることで、現象を捉える視点は飛躍的に複雑さを増す。こうした視点の複雑さは、研究デザインやデータ分析の方法論の複雑化という代償とともに、横断研究とは全く異なる種類の創造的なりサーチクエスションの創出をもたらすことで、発達研究の広がりや深まりに貢献するであろう。

文 献

- Adachi P. & Willoughby T. (2015). Interpreting effect sizes when controlling for stability effects in longitudinal autoregressive models: implications for psychological science. *European Journal Developmental Psychology*, *12*, 116-128.
- Angrist, J.D., Imbens, G.W., & Rubin, D.B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, *91*, 444-455.
- Bailey, D.H., Oh, Y., Farkas, G., Morgan, P., & Hillemeier, M. (2020). Reciprocal effects of reading and mathematics? Beyond the cross-lagged panel model. *Developmental Psychology*, *56*, 912.
- Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, *3*, 133-153.
- Cattan, S., Kamhöfer, D.A., Karlsson, M., & Nilsson, T. (2022). *The long-term effects of student absence: Evidence from Sweden* (No. 383). DICE Discussion Paper.
- Cook, T.D., Campbell, D.T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Coon, D., & Mitterer, J. (2013). *Psychology: A journey* (5th ed.). Belmont, CA: Wadsworth Cengage Learning.
- Erikson, E.H. (1959). *Identity and the life cycle*. WW Norton & Company.
- Glenn, N.D. (2005). *Cohort analysis* (2nd ed.). London: Sage.
- Hamaker, E.L., Kuiper, R.M., & Grasman, R.P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*, 102.
- 伊藤大幸・浜田 恵・村山恭朗・高柳伸哉・野村和代・明畷光宜・辻井正次. (2017). クラスサイズと学業成績および情緒的・行動的問題の因果関係: 自然実験デザインとマルチレベルモデルによる検証. *教育心理学研究*, *65*, 451-465.
- 伊藤大幸・浜田 恵・村山恭朗・高柳伸哉・明畷光宜・辻井正次. (2021). 小中学生の自由時間の活動が心理社会的適応に及ぼす影響に関する縦断的検証. *発達心理学研究*, *32*, 91-104.
- 児玉(渡邊)茉奈美・榎原良太. (2022). 子どもへの叱りと否定的感情の関連およびその調整要因の検討: 経験サンプリング法を用いた4週間の縦断調査. *発達心理学研究*, *33*, 356-365.
- Lucas, R.E. (2022). It's time to abandon the cross-lagged panel model. <https://psyarxiv.com/pkec7> PsyArxiv Preprint. (2022年8月29日10時00分)
- Lüdtke, O., & Robitzsch, A. (2021). A critique of the random intercept cross-lagged panel model. <https://psyarxiv.com/6f85c> PsyArxiv Preprint. (2022年8月29日10時00分)
- 西村倫子. (2022). 発達研究における縦断データの解析手法: 成長曲線モデルと潜在クラス成長分析. *発達心理学研究*, *33*, 256-266.
- Orth, U., Clark, D.A., Donnellan, M.B., & Robins, R.W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*, 1013.
- Piaget, J. (1970). Genetic epistemology. In *Genetic epistemology*. NY: Columbia University Press.

- Platt, B., Kadosh, K. C., & Lau, J. Y. (2013). The role of peer rejection in adolescent depression. *Depression and Anxiety*, *30*, 809–821.
- Mason, K. O., Mason, W. M., Winsborough, H. H., & Poole, K. (1973). Some methodological issues in cohort analysis of archival data. *American Sociological Review*, *38*, 242–258.
- McGuire, T. C., McCormick, K. C., Koch, M. K., & Mendle, J. (2019). Pubertal maturation and trajectories of depression during early adolescence. *Frontiers in Psychology*, *10*, 1362.
- Murayama, Y., Ito, H., Hamada, M., Takayanagi, N., Myogan, M., Suzuki, K., & Tsujii, M. (2021). Examining simultaneous associations of four emotion regulation strategies with abnormal eating behaviors/attitudes in early adolescents. *Eating Behaviors*, *40*, 101449.
- Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*, 463–469.
- Katagiri, M., Ito, H., Murayama, Y., Hamada, M., Nakajima, S., Takayanagi, N., Uemiya, A., Myogan, M., Nakai, A., & Tsujii, M. (2021). Fine and gross motor skills predict later psychosocial maladaptation and academic achievement. *Brain and Development*, *43*, 605–615.
- Rosenbaum, P. R. & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* *70* (1), 41–55.
- Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, *51*, 309.
- 宇佐美慧. (2022). 個人内関係の推測と統計モデル：ランダム切片交差遅延パネルモデルを巡って. *発達心理学研究*, *33*, 267–286.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MA: MIT.
- Yang, Y., & Land, K. C. (2013). *Age-period-cohort analysis: New models, methods, and empirical applications*. Boca Raton: CRC.
- Yao, Q., Wang, P., Wang, X., Qie, G., Meng, M., Tong, X., Bai, X., Ding, M., Liu, K., & Chu, Y. (2020). Retrospective study of risk factors for severe SARS-Cov-2 infections in hospitalized adult patients. *Pol Arch Intern Med*, *130* (5), 390–399.
- 山田剛史. (2022). 一事例実験のデザインと分析方法について. *発達心理学研究*, *33*, 287–303.

Ito, Hiroyuki (Human Science Division, Faculty of Core Research, Ochanomizu University). *Role and Methodology of the Longitudinal Approach in Developmental Research: What, Why, and How?* THE JAPANESE JOURNAL OF DEVELOPMENTAL PSYCHOLOGY 2022, Vol.33, No.4, 176–192.

In this paper, we discuss what longitudinal research is, why it is necessary, and how to conduct it in real developmental-research context. Longitudinal research can be conducted using various research designs, but their common features are that data are collected from the same subject at multiple time points and analyses focus on change over time and relationships within individuals. Longitudinal approaches with these characteristics are beneficial to developmental research because they enable the researcher to (1) separate the effects of age, cohort, and period, (2) quantify trajectories of intraindividual change and how these differ between individuals, and (3) identify clues to causal relationships (especially temporal order). Taking a longitudinal approach increases the complexity of research design and data-analysis methodology. However, this cost is offset by the potential to broaden and deepen developmental research by generating creative research questions of a different kind from those generated by cross-sectional studies.

[Keywords] Longitudinal research, Age-period-cohort analysis, Developmental trajectories, Causal inference

2022.5.17 受稿, 2022.8.29 受理