

## ЛОНГИТЮДНЫЕ ИССЛЕДОВАНИЯ: ТЕОРИЯ И МЕТОДЫ

*КОРНИЛОВ С. А., Московский государственный университет им. М. В. Ломоносова, Москва*

В статье рассматривается современное понимание лонгитюда как группы исследовательских методов, позволяющих проверять особый класс гипотез (о развитии, его форме и динамических характеристиках). Дается описание основных лонгитюдных схем и их возможных ограничений, а также проводится анализ угроз валидности лонгитюдных исследований. Статья завершается кратким описанием основных статистических подходов к анализу лонгитюдных данных.

**Ключевые слова:** лонгитюд, лонгитюдное исследование, повторяющиеся измерения, лонгитюдная валидность, иерархическое линейное моделирование.

В психологии лонгитюдный метод обычно противопоставлялся методу срезов, а в последнее время стал также рассматриваться в контексте поиска отсроченных экспериментальных эффектов. Другое его название – это метод «продольных срезов», противопоставлявшийся методу «поперечных срезов». Если метод срезов (или кросс-секционный метод) предполагает сопоставление психологических показателей в одно и то же время (их измерения) у людей разных возрастных групп, то классический лонгитюд означает «продолженное исследование» – с многократной фиксацией показателей на одном и том же человеке или одной и той же группе (выборке) людей.

Лонгитюдные исследования занимают особое место в методологии психологических исследований, а также социальных наук и, как указывается обычно в зарубежной литературе, наук о поведении. Это связано как со спецификой проверяемых в лонгитюде гипотез о развитии, так и со сложностями планирования, проведения и обработки данных лонгитюдных исследований. Нам представляется важным отметить, что слово «развитие» здесь понимается нами в самом общем виде как процесс изменения (англ. *change*), – развитие, не несущее само по себе психологического содержания. Гипотезы о развитии включают предположение о временной динамике изменений психологических показателей изучаемого базисного процесса. Фактор времени при этом не рассматривается как источник или причинно действующий фактор развития, он выступает лишь аналогом независимой переменной (НП). Теоретическое обоснование того, может ли временная динамика изменений психологических показателей трактоваться как развитие, включает и методологические принципы понимания развития, и положения конкретной психологической теории, и оценку планирования исследования (с целью организации достоверных выводов).

Вместе с тем, лонгитюдные исследования позволяют прямо обратиться к проверке каузальных гипотез со стороны требования к временной последовательности причин и следствий и, таким образом, могут приближаться к выполнению двух основных условий установления каузальной связи – следования причины и следствия во времени и ковариации между ними (Корнилова, 2003). Место причинных при этом могут занимать любые наблюдаемые воздействия, которые, однако, не могут трактоваться в качестве экспериментальных, если психолог не управляет ими. Два других требования причинного вывода могут быть удовлетворены в срезовых или последовательных кросс-секционных исследовани-

ях. Так, требование к наличию ковариации между переменными устанавливается либо через межгрупповые различия, либо ненулевые корреляции между переменными, а требование к отсутствию альтернативных объяснений может быть удовлетворено путем использования средств экспериментального или статистического контроля.

Методы лонгитюдных исследований, возникновение которых исторически связывается с началом ведения систематических переписей населения в Канаде и Квебеке в XVII веке, получили наибольшее развитие после Первой мировой войны в США, а затем в конце XX века в социальных науках и науках о поведении (Menard, 2002). Их современное развитие прямо связывается с развитием соответствующих методов анализа данных, определяемых на стадии планирования исследования. Недавно опубликованную в журнале *Journal of Management* статью, посвященную лонгитюдным методам в науках об организационном поведении, Р. Плойхарт и Р. Ванденберг начинают с указания на тот факт, что в большинстве современных теорий прямо или косвенно делаются утверждения, имеющие динамический характер, т.е. апеллирующие к объяснению какого-либо феномена в контексте его динамики или динамики его связей с другими феноменами (Ployhart, Vandenberg, 2010). То же самое можно сказать и о психологических закономерностях, устанавливаемых в контексте проверки гипотез о развитии, а также о длительных или отсроченных эффектах экспериментальных воздействий.

Анализ литературы, посвященной планированию психологических исследований, выявляет отсутствие четкого определения лонгитюда. В самом общем виде лонгитюд можно понимать как *группу методов* (Wall, Williams, 1970), характеризующихся, в противовес срезовым исследованиям, наличием нескольких повторяющихся измерений одной или более интересующих исследователя переменных на материале одних и тех же или схожих групп испытуемых.

В качестве наиболее известных и получивших социальное звучание можно привести примеры: 1) лонгитюда Л. Термена на выборке более 1000 интеллектуально одаренных детей, начатого в 1921 году и продолжающегося до сих пор (Terman, Oden, 1947); 2) геронтологического лонгитюда под руководством Г. Томэ (Thomae, 1993), начатого после Второй мировой войны на выборке более 10 000 человек в Германии и продолжавшегося 20 лет, который показал первенствующую роль самосознания личности в динамике развития личностно-мотивационной сферы. Как и в случае первого примера, выводы имели широкую сферу обобщений. Этому способствовали указанные величины выборок и количество измерений, но также и такой аспект организации исследования, как проведение многократного тестирования одних и тех же испытуемых, что позволяло включать в интерпретацию динамики показателей предположения о развитии.

Несмотря на то, что проверка гипотез о развитии – несомненная прерогатива лонгитюдных исследований, выводы о развитии часто совершаются исходя из результатов эмпирических исследований, выполняемых в рамках различных психологических теорий при использовании срезового метода. Последний позволяет выявлять связи между несколькими «статичными» переменными, взятыми в отдельный период времени; основанием для указанной претензии на выводы о развитии служит негласное предположение об эквивалентности выборок («сквозь» которые проводится сравнение) и исторических периодов для разных когорт испытуемых, что нередко приводит к игнорированию важного для лонгитюда источника смещений, который заслуживает особого внимания.

## Ключевые понятия лонгитюдного метода

Термин «*когорта*» чаще всего используется для обозначения общности людей в выборке по критерию года рождения (*birth cohort*) и, согласно его демографическому определению (см., например: Glenn, 1977), означает группу людей в рамках географически или иначе обозначенной популяции, пережившую схожие события в указанный период времени. Переменная *возраста* – это хронологический возраст испытуемых в момент измерения показателей. Термин «*период*» требует дополнительных уточнений – им обозначают и время измерения, скажем, календарный год, и период, охватываемый временем жизни когорты, включающий общую для нее историю. Формально когорта в общем виде определяется как:

Когорта (год рождения) = Период измерения (календарный год) – Возраст (количество лет с момента рождения).

Эта простая формула иллюстрирует линейную зависимость когорты, времени измерения (которому сопутствует общая история периода) и возраста при проведении лонгитюда, эксплицируя самый важный для лонгитюда источник *систематических смещений*. Люди одного года рождения живут в общих для них социальных условиях развития, выпадающих на определенный исторический период. Тем самым общим для такой когорты оказывается не только год рождения, но и «история» – заполнение соответствующего временного промежутка общностью социума, в котором они живут (в определенной стране, в географически, экономически, политически и культурно ограниченном пространстве). Игнорирование этого смещения ставит под угрозу валидность выводов лонгитюдного исследования, что будет проиллюстрировано ниже.

Итак, указанная линейная зависимость эффектов когорты, возраста и времени измерения ведет к тому, что при контроле любых двух переменных из указанных трех осуществляется контроль и третьей. Если исследование проводится методом срезов, то выборка людей в каждом определенном срезе имела также одну «историю», но эти «истории» различались для участвующих в лонгитюде секций и срезов, что привело к смещению возраста и фактора социальных условий (периода). Поэтому при проведении срезовых сравнений психологических показателей у людей разного возраста обнаруженные различия между более молодыми и более зрелыми испытуемыми могут отражать не линию развития интересующего исследователя базисного процесса, но эффекты когорты. Проведение же лонгитюдного исследования с последовательными множественными измерениями указанного базисного процесса на материале единственной когорты может привести к выявлению не выделенных в качестве предмета изучения эффектов развития, а эффектов социальных условий как исторического периода, специфичного для данной когорты.

Попытки преодоления линейной зависимости трех типов эффектов разделяются на две концептуальные группы. Первая, представленная исследованиями К. Мэйсона (Mason et al., 1973), представляет собой попытку решения проблемы на статистическом уровне путем создания моделей, устраняющих статистическую *коллинеарность* (т. е. полную математическую зависимость) между возрастом, когортой и временным периодом. Вторая группа представляет собой подходы, предполагающие теоретическое обоснование исключения рассмотрения влияния одной из трех переменных на обнаруживаемые линии развития или их переосмысления. В этом русле были разработаны подходы, рассматривающие эффекты когорты как взаимодействие эффектов возраста и периода (Menard, 2002) или заменяющие

когорту ее характеристиками, поддающимися точному определению и измерению, например, сравнительным размером когорты (Ryder, 1965). В идеальном случае эффекты *периода* и *когорты*, имеющие в методологии лонгитюдных исследований принципиально отличный от эффектов времени объяснительный статус, исключаются из каузального анализа. Они заменяются операционализированными характеристиками, позволяющими проводить точный контроль эффектов этих переменных, что позволяет разводить эффекты возраста (или эффекты развития), эффекты периода (исторические эффекты) и эффекты когорты. Такой анализ принципиально невозможен вне рамок «истинного», или панельного (см. ниже), лонгитюдного исследования, в котором множество измерений совершается на материале нескольких когорт сразу (см. рис. 1; лонгитюдно-последовательная стратегия). Основные схемы лонгитюда рассмотрены Г. В. Бурменской (2004) в учебном пособии «Эксперимент и квазиэксперимент в психологии» под ред. Т. В. Корниловой и представлены в приводимых ею схемах К. Шайе (рис. 1). В верхней части рисунка по горизонтали фиксируется (выделена длинным прямоугольником) схема лонгитюда, по диагонали (в один и тот же момент измерения) – метод срезов, имеющий название также сравнительно-возрастного, а по вертикали прочитывается схема для метода временного запаздывания, о которой мы скажем позднее. Вторая часть (снизу) представляет современные подходы к планированию лонгитюда.

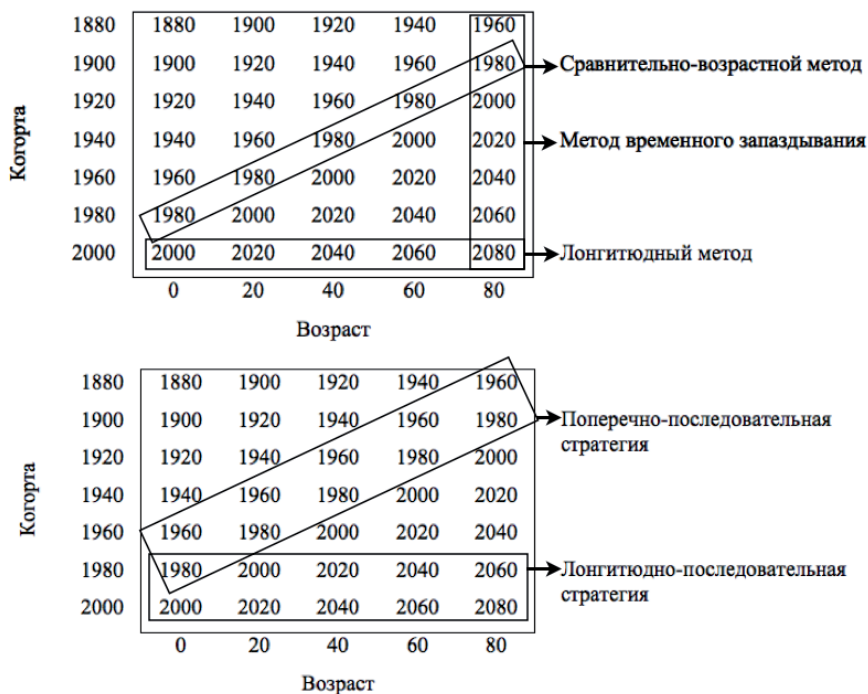


Рис. 1. Классическая и модифицированная схемы К. Шайе, приводимые Г. В. Бурменской (2004)

Представленные на рис. 1 схемы дают представление о разных стратегиях сбора данных в умозрительном двухсотлетнем исследовании, охватывающем семь когорт. Здесь нет предположений об экспериментальных влияниях. Отметим, тем не менее, что за рамками возрастной психологии лонгитюд используется и при проверке каузальных гипотез; в этом случае осуществ-

вляется контроль уровней НП, а лонгитюдные изменения ЗП высвечивают временную динамику изменений в базисном процессе, вызванную именно экспериментальным воздействием.

### Эффекты, фиксируемые в лонгитюдных исследованиях

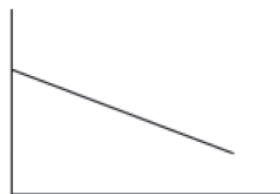
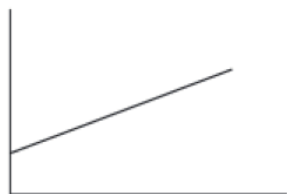
Остановимся на такой особенности современных лонгитюдных исследований, как возможность проверки «сильных» каузальных гипотез при количественной оценке динамических характеристик развития. Основными целями проведения лонгитюдных исследований (Cook, Ware, 1983; Menard, 2002; Singer, Willet, 2003; Ployhart, Vandenberg, 2010) считаются:

- увеличение *точности измерения экспериментального эффекта* путем контроля интраиндивидуальной изменчивости. В этом случае применяются интраиндивидуальные схемы исследования, или схемы *повторяющихся измерений (repeated measures design)*, к которым и относится группа лонгитюдных методов, включая схему тестирования «до–после»;
- проверка гипотез о направленности каузальных связей и оценка их силы;
- установление функциональной формы *интраиндивидуальных траекторий*, или кривых развития;
- оценка *межиндивидуальных различий в интраиндивидуальных траекториях развития* путем построения каузальных моделей, их предсказывающих.

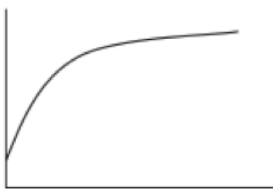
Ключевые различия в понимании лонгитюда в литературе связаны с отсутствием консенсуса в отношении минимального количества временных срезов. С. Менард (Menard, 2002) считает, что любое исследование с двумя и более повторяющимися измерениями можно считать лонгитюдным, тогда как Р. Плойхарт и Р. Ванденберг (Ployhart, Vandenberg, 2010), а также Дж. Зингер и Дж. Уиллет (Singer, Willet, 2003) указывают на принципиальные ограничения исследований с двумя «волнами» сбора данных (*data collection waves*), поскольку такие исследования не позволяют оценить *функциональную форму* развития, а также подвержены влияниям со стороны случайных *ошибок измерения*, которые при наличии всего двух временных срезов способны привести исследователей к неверным заключениям о наличии и/или направленности эффекта.

Другое различие в понимании лонгитюда связано со стремлением некоторых авторов (см., например: Baltes, Nesselroade, 1979; Wall, William, 1970) исключить из лонгитюдных схем все, кроме схемы *проспективного панельного лонгитюда (prospective panel design)*. Проспективные панельные дизайны характеризуются тем, что сбор данных в отношении интересующих исследователя временных отрезков проводится в указанные отрезки, т.е. сбор данных проводится «здесь и сейчас» для каждого временного отрезка. *Ретроспективные панельные дизайны (retrospective panel design)* предполагают, что сбор данных в отношении различных временных периодов в прошлом может быть осуществлен «здесь и сейчас», т.е. в проспективной схеме данные собираются в соответствующее каждому отрезку время, тогда как ретроспективная схема позволяет собирать данные о событиях в прошлом испытуемых. Тем не менее, в обоих случаях данные собираются для одних и тех же выборок испытуемых. В отличие от указанных двух схем срезовой метод, также получивший название *повторяющегося кросс-секционного (repeated cross-sectional)* лонгитюда, рассматривает временные периоды как отдельные срезы, собранные на разных испытуемых, но признание выборок взаимозаменяемыми (что требует специального планирования, обоснования и проверки) предположительно обеспечивает валидность указанных кросс-секционных, или срезовых, сравнений.

## Линейные формы развития



## Нелинейные формы развития



## Прерывающиеся формы развития

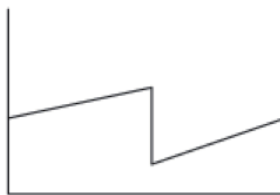
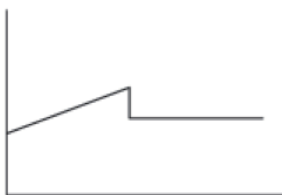


Рис. 2. Различные формы развития, обнаруживаемые в лонгитюдных исследованиях. По горизонтали на всех графиках откладывается время

### Гипотезы, проверяемые в лонгитюдном исследовании

Как уже было отмечено выше, проведение лонгитюдного исследования преимущественно связано с проверкой гипотез о развитии, или гипотез о наличии интраиндивидуальной изменчивости в показателях НП с течением времени. Фактор времени при этом полагается фундаментальным предиктором траекторий развития, но не экспериментальным воздействием.

Формулировка гипотез об *интраиндивидуальной изменчивости* включает утверждения о функциональной форме кривых развития, задаваемой на основе теоретических положений в отношении конкретного феномена, а также на основе первичного (в том числе визуального) анализа данных. При этом, как показано на рис. 2, рост может быть линейным и нелинейным, прерванным (в контексте изменения свойства, например, в связи с осуществлением экспериментального воздействия в середине множественных измерений), положительным, отрицательным и т. д. Уточнение функциональной формы кривых развития критично для фальсификации гипотез, проверяемых в лонгитюде (Mitchell, James, 2001). Обнаруживаемые кривые развития позволяют как отклонять предварительные гипотезы (в случае несоответствия им полученных лонгитюдных данных), так и формулировать новые. В этом аспекте гипотетико-дедуктивное рассуждение сближает метод лонгитюда с экспериментальным.

Лонгитюд также позволяет проверять гипотезы о наличии *межиндивидуальных различий* в интраиндивидуальных кривых развития, т. е. построения моделей, предсказываю-



щих индивидуальные траектории развития в зависимости от набора константных или варьирующих во времени групповых и индивидуальных характеристик. Последнее существенно для современных лонгитюдных исследований в связи с развитием представлений об индивидуальном характере каузальных эффектов (Raudenbush, 2001), т. е. пониманием эффектов воздействия НП не как константных, а как зависящих от того или иного набора индивидуальных различий в биологических, социальных и психологических факторах и свойствах человека. Этот подход, названный подходом Рубина-Розенбаума-Холланда (RRH: Rubin, 1974; Rosenbaum, Rubin, 1983; Holland, 1986), позволяет моделировать величину экспериментального эффекта в зависимости от набора измеряемых переменных, которые выступают предикторами, т. е. аналогами НП, по отношению уже к самому показателю экспериментального эффекта (выступающему в современных вариантах развития методов регрессионного анализа в качестве ЗП).

Помимо указанных выше гипотез, лонгитюдные исследования позволяют проверять и другие. Так, целью лонгитюда может быть установление *исторических трендов* в развитии какого-либо свойства (см., например, исследования развития уровня интеллекта среди различных наций; Flynn, 1996) или даже временных или исторических трендов *в развитии паттернов взаимосвязей* между переменными, измеренных неоднократно: С. Менард приводит пример своего анализа лонгитюдных данных крупного проекта National Youth Survey (Menard et al., 1989), показавшего, что вступление в брак в подростковом возрасте положительно связано с использованием запрещенных веществ и психиатрическими расстройствами, но вступление в брак в период ранней взрослости отрицательно связано с этими же самыми показателями (см.: Menard, 2002).

Другой важной областью использования лонгитюдного метода является включенность его в схемы экспериментальной проверки каузальных гипотез. Во-первых, как уже было отмечено, использование лонгитюдного метода в целях получения нескольких волн данных «до» и «после» экспериментального воздействия в рандомизированных экспериментах позволяет значительно увеличить мощность исследования путем контроля межгрупповых различий в начальных показателях ЗП (использование лонгитюда в экспериментальных исследованиях подробно описано: Twisk, 2008). Во-вторых, применение метода запаздывания, отсроченных кросс-корреляций или линейного панельного анализа (*linear panel analysis*) позволяет относительно консервативно (Menard, 2008) проверять гипотезы о наличном и отсроченном влиянии одних переменных на другие при невозможности проведения рандомизированного эксперимента. *Панельный анализ* предполагает построение модели, в которой предшествующие значения переменных предсказывают как их будущие значения, так и значения других переменных после вычета вклада их собственных предшествующих значений. Лонгитюдный метод отличается и возможностью проверки так называемых рекурсивных каузальных моделей, в которых несколько переменных взаимовлияют друг на друга.

### Основные проспективные лонгитюдные схемы

Существующие разногласия в понимании типов исследований, которые можно считать лонгитюдными, описанные выше, разумеется, касаются и проблемы классификации конкретных лонгитюдных схем. Остановимся на классификации, предложенной С. Менардом (Menard, 2002, 2008) поскольку она получила наиболее широкое распространение в зарубежной литературе по методологии лонгитюдных исследований. Согласно С. Менарду,

существует четыре принципиальные схемы лонгитюдного исследования, представленные на рис. 3, где по горизонтали отмечаются временные периоды, а по вертикали – выборки лиц, участвующих в исследовании.

Первая схема – схема полного популяционного дизайна (*total population design*) – предполагает обследование всей интересующей исследователя популяции (например, всех жителей административного округа) на протяжении всего целевого временного периода. Поскольку от периода к периоду происходит изменение состава популяции (определяемое через количество смертей и новых рождений), выборки оказываются неидентичными для различных временных отрезков. Такая схема исследования позволяет проводить оценку *временных трендов*, но представляет определенную трудность для исследователя, поскольку обнаруженная динамика может быть связана с изменением состава популяции. Использование полного популяционного дизайна позволяет проводить кросс-секционные сравнения между включенными в популяцию когортами, а также отслеживать динамику для отдельных когорт. Это одна из двух схем, позволяющих проводить все типы анализа лонгитюдных данных.

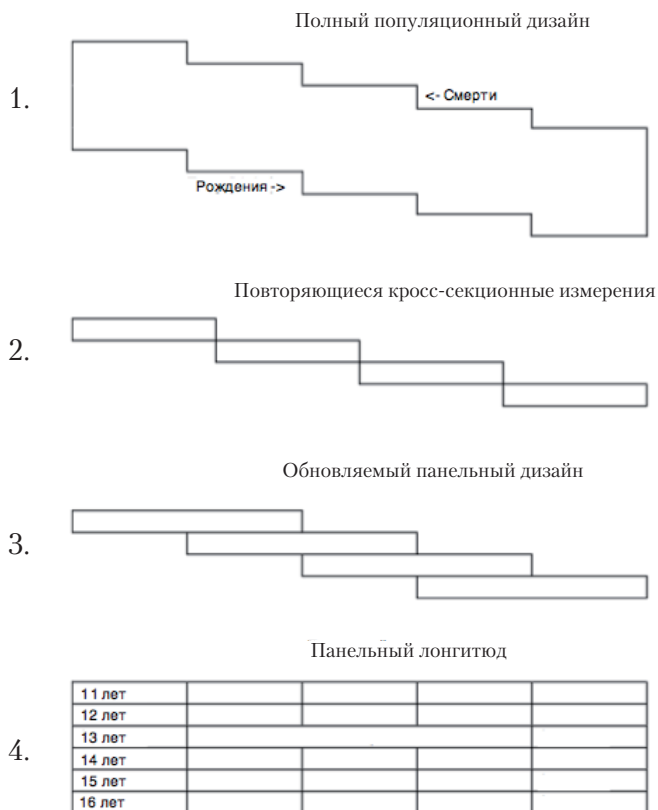


Рис. 3. Основные проспективные схемы лонгитюдных исследований. По горизонтали откладываются временные периоды, по вертикали – выборки, для которых проводится сбор данных в указанный временной период



Вторая схема – повторяющихся срезовых, или кросс-секционных, измерений (*repeated cross-sectional design*) – включает измерение показателей на материале случайных выборок, отличающихся для каждого временного периода. Даже при выполнении требования о сопоставимости (эквивалентности) выборка эта схема не позволяет проводить ни тонкий анализ паттернов развития для отдельных когорт, ни процедуры уточнения каузальных связей между переменными. Поэтому она чаще всего используется для изучения эффектов исторического периода и для множественной репликации результатов предыдущих кросс-секционных исследований.

Третья схема – обновляемого панельного лонгитюда (*revolving panel design*). Ее применение позволяет исследователю проводить множественные измерения на определенной выборке испытуемых в течение определенного промежутка (или промежутков) времени, а затем заменять часть выборки или всех испытуемых новыми, что позволяет преодолеть ограничения, связанные с «выпадением» испытуемых из исследования, а также увеличить надежность данных при ретроспективном сборе данных (т.е. сборе данных о периодах в прошлом). Такая схема позволяет проводить анализ кратковременных паттернов изменений на уровне отдельных испытуемых и когорты, а также панельный анализ. Применение схемы обновляемого панельного лонгитюда особенно плодотворно при изучении специфических групп с жестко заданными возрастными ограничениями.

Четвертая схема – «истинно» лонгитюдная схема, или схема панельного лонгитюда (*longitudinal panel design*). Она предполагает сбор данных для всех временных периодов на материале множества когорт. Главными достоинствами этой схемы, выделяющими ее из группы лонгитюдных методов, являются возможность проведения любого типа анализа лонгитюдных данных и точная квантификация и разведение эффектов когорты и времени.

### **Угрозы валидности лонгитюдных исследований**

Помимо указанных выше проблем, связанных со смешениями при организации лонгитюдных исследований, их проведение включает также контроль как типичных угроз валидности, так и особого класса угроз внутренней валидности, характерных практически исключительно для лонгитюдных исследований. Так, проведение лонгитюдного исследования предполагает планирование количества временных срезов и промежутков между ними. Ключевой направляющей такого планирования является предполагаемая форма развития, на выявление которой направлен лонгитюд и которая накладывает ограничения на минимальное количество срезов (увеличивающееся при усложнении гипотетической формы развития). Увеличение количества срезов ведет к повышению надежности данных и увеличению статистической мощности (Willet, 1989), но соображения практического характера ограничивают возможное количество срезов в любом лонгитюдном исследовании. Недостаточное количество временных срезов или слишком длинные промежутки между ними могут привести к тому, что полученные данные будут неадекватно описывать выявляемые паттерны изменений. К таким искажениям могут также привести либо неадекватный учет предполагаемой формы кривых развития (например, в случае их нелинейности), либо игнорирование временных параметров этих изменений (темпа). Отметим, что при наличии адекватно выбранных временных промежутков для измерения психологических переменных срезы показателей в лонгитюдном исследовании могут проводиться не только через равные промежутки времени, но и неравномерно (например, если наиболее тонкие механизмы развития предполагаются в начале исследования, количество срезов в этот период можно увеличить).

Другой выступает угроза *лонгитюдной валидности* (*longitudinal validity*), понимаемой как инвариантность измеряемых конструкторов на протяжении проведения лонгитюда (Vandenberg, 2002). Низкая лонгитюдная валидность может быть связана как с техническими проблемами (например, смена лаборантского состава), так и с концептуальными. Во втором случае перед проведением анализа лонгитюдных данных исследователю необходимо проверить соответствие их критерию *структурной валидности* (взаимосоответствие измеряемых в разные временные отрезки времени конструкторов, например, идентичность факторной структуры) путем проверки гипотезы об отсутствии инвариантности операционализируемых конструкторов на всем протяжении лонгитюда. В случае неподтверждения гипотезы об отсутствии конструктивной инвариантности исследователь переходит к следующему этапу – проверке идентичности метрики используемых шкал. Такой тест, например, возможно осуществить путем сопоставления факторных нагрузок пунктов опросников, взятых на различных временных срезах. Проблема смещений в структуре операционализируемых конструкторов или в метрике их измерения требует особого учета на этапах планирования лонгитюдного исследования и первичного анализа лонгитюдных данных (Singer, Willet, 2003). Валидность выводов лонгитюдного исследования, таким образом, зависит от того, насколько изменялась структура изучаемого свойства на протяжении исследования и насколько это отражено в изменениях в структуре методик, используемых для операционализации переменных.

Еще одной угрозой валидности лонгитюдного исследования является *потеря испытуемых* (их «выпадение»), или «истощение» (*attrition*), выборки на протяжении исследования. Практически все лонгитюдные исследования сталкиваются с этой проблемой: в отдельных случаях выборка к концу лонгитюда сокращается больше чем наполовину (например, в исследованиях использования подростками запрещенных веществ: Newcomb, Bentler, 1988; см.: Menard, 2002). Критическим для лонгитюда является вопрос, может ли такое «истощение» выборки привести к систематическим смещениям. Ответ на этот вопрос частично зависит от того, отличаются ли систематически испытуемые, оставшиеся в лонгитюде, от тех, кто вышел из исследования, по интересующим исследователя характеристикам. Эта проблема прямо связана с проблемой *пропущенных данных* (*missing data*), требующей от исследователя изучения причины потери испытуемых.

Вслед за Рубином (Rubin, 1976) большинство исследователей разделяют *совершенно случайно пропущенные данные* (*missing completely at random, MCAR*) и *случайно пропущенные данные* (*missing at random, MAR*). MCAR данными признаются пропущенные данные, которые не связаны с переменными в исследовании. MAR данные – данные, которые не связаны с переменными в исследовании после проведения контроля релевантных переменных. Если пропущенные данные нельзя считать ни MCAR, ни MAR, они признаются *не случайно пропущенными* (*missing not at random, MNAR*). MCAR и MAR данные могут быть «включены» в лонгитюдный анализ, тогда как игнорирование MNAR данных является игнорированием важного источника систематических смещений. Методологи лонгитюда (Graham, 2009; Graham, Hofer, 2000; Singer, Willet, 2003) предлагают различные альтернативы анализа лонгитюдных данных с пропущенными значениями (преимущественно через использование сложных современных статистических моделей вычисления, например, метода полного максимального подобия, FIML – *full information maximum likelihood*) или ввода данных (например, MI – *multiple imputation*, множественный ввод), но сходятся во мнении относительно недопустимости их игнорирования.

Предотвращение появления (или хотя бы снижение количества) пропущенных данных требует от исследователя как теоретического анализа потенциальных источников пропущенных данных на этапе планирования исследования, так и разработки стратегий поддержания контакта с испытуемыми. Такие стратегии включают получение контактной информации родственников и друзей испытуемых, когда это возможно, поддержание контакта путем отправки поздравительных открыток, а также использование телефонных и адресных справочников и в последнее десятилетие – ресурсов интернета.

### Анализ лонгитюдных данных

Сложность, характеризующая лонгитюдный метод как группу исследовательских схем, в полной мере распространяется на статистический анализ получаемых лонгитюдных данных. Нормативы такого анализа претерпели существенные изменения в последние десятилетия, существенно расширив доступные психологу альтернативы. Выбор аналитических стратегий зависит как от лонгитюдного дизайна (количества временных срезов и их временной удаленности) и проверяемых в нем гипотез, так и от уровня статистической подготовки самого исследователя.

С. Менард (Menard, 2002, 2008) иллюстрирует зависимость аналитических стратегий от дизайна исследования и проверяемых гипотез, проводя различия между тремя основными областями применения лонгитюда, из которых нас особо интересуют две основные: первая – экспериментальные и квазиэкспериментальные исследования; вторая – исследования роста и развития. В первом случае исследователя интересуют преимущественно кратковременные изменения и различия в них между экспериментальной и контрольной группами после осуществления воздействия. Во втором случае выявляются долгосрочные паттерны изменений и их систематические связи с различными характеристиками испытуемых. Дизайн исследования включает планирование размера выборки, и два указанных типа исследований различаются по этому параметру (большие  $n$  в первом и меньшие – во втором случае), что также регулирует выбор методов анализа. Так, для больших выборок и большого количества срезов наиболее информативны *многомерные модели кривых развития (multilevel growth curve models)*, а для маленьких выборок и большого количества срезов – ауторегрессионный анализ *временных сечений (autoregressive time series models)*. Дж. Зингер и Дж. Уиллет (Singer, Willet, 2003), а также С. Менард (Menard, 2002, 2008) дают подробное описание указанных и других методов анализа, мы же остановимся на критических различиях между «классическим» для отечественных лонгитюдных исследований подходом – использованием ANOVA для *повторяющихся измерений (repeated measures ANOVA)* – и современными методами, включая *иерархическое линейное моделирование (hierarchical linear modeling, HLM; аналогично – multilevel growth curve modeling)* и моделирование изменений *латентных переменных в рамках структурного моделирования*, или моделирования на основе структурных уравнений (*latent growth curve modeling; structural equation modeling, SEM*).

Наиболее часто используемый метод анализа лонгитюдных данных, ANOVA для повторяющихся измерений, входит в класс *общих линейных моделей (general linear models, GLM)*, основной целью которых является выявление межгрупповых различий в средних значениях изменяющихся со временем показателей. Использование GLM в лонгитюдных исследованиях ограничивается частотой нарушений статистических требований GLM, как, например, независимость и гомогенность остатков (так называемая проблема *сферичности – sphericity*), что ведет к искажениям при проверке статистических гипотез с помощью этого

метода. Другим ограничением этого метода является *невозможность* квантификации интраиндивидуальных различий в показателях изменений, неприспособленность этого метода к использованию недихотомических факторов и предикторов, сложности в преодолении проблемы пропущенных данных и невозможность построения истинно нелинейных моделей (включающих экспоненты), а также использования варьирующих во времени предикторов.

Большая часть из этих ограничений снимается при использовании двух связанных групп методов, являющихся современным вариантом развития регрессионного анализа, а именно методов иерархического линейного моделирования (*hierarchical linear modeling*, HLM) и структурного моделирования (*structural equation modeling*, SEM). Оба класса методов позволяют моделировать сложные кривые развития путем установления соответствия данных гипотетическим моделям роста. Кратко опишем эти методы.

HLM, также получивший название многоуровневого анализа кривых роста (*multilevel growth curve modeling*), предполагает построение многоуровневых моделей изменений, при этом первый уровень является уровнем *множественных измерений*, а второй – уровнем *индивидуальных различий* (Raudenbush, Bryk, 2002), что соответствует основным указанным гипотезам, проверяемым в лонгитюде:

$$\text{Level1} : Y_{it} = \pi_{0i} + \pi_{1i}T_{it} + e_{it}$$

$$\text{Level2} : \pi_{0i} = \beta_{00} + \beta_{01}X_i + r_{0i}$$

$$\text{Level2} : \pi_{1i} = \beta_{10} + \beta_{11}X_i + r_{1i}$$

Отметим, что все три приведенных выше уравнения можно заменить одним путем подставления значений в Level 1. При этом общая модель может быть условно подразделена на структурный и стохастический компоненты. Если  $Y_{it}$  – значение по зависимой переменной для испытуемого  $i$  на временном этапе  $t$ , то структурные компоненты (фиксированные эффекты, *fixed effects*) –  $\pi_{0i}$  – так называемый начальный статус (*intercept*) при  $t = 0$ ,  $\pi_{1i}$  – показатель прироста в значениях по зависимой переменной, соответствующего изменению в 1  $T$  (*slope*). Стохастические компоненты (случайные эффекты, *random effects*) –  $e_{it}$ ,  $r_{0i}$ ,  $r_{1i}$ , отражающие индивидуальные эффекты, ошибку измерения и необъясненную дисперсию.

Модель на первом уровне (Level 1) описывает интраиндивидуальные изменения в показателях интересующей исследователя переменной в течение времени ( $T$ ; каждый испытуемый может иметь свою индивидуальную траекторию изменения); модель на втором уровне (Level 2) задает межиндивидуальные различия в темпах изменений. При этом начальный уровень ( $\pi_{0i}$ ) и динамические показатели роста ( $\pi_{1i}$ ) могут быть смоделированы в терминах предикторов (как дихотомических – принадлежность к группе, так и количественных). Таким образом, путем концептуального переосмысления и математического переопределения единиц анализа лонгитюдных данных через сведение их в иерархическую систему HLM позволяет оценивать и моделировать как интраиндивидуальные различия, так и межиндивидуальные различия в кривых развития в лонгитюдном исследовании, с легкостью позволяет использовать пропущенные значения (в рамках использования метода полного максимального подобия) и проверять сложные гипотезы о нелинейных изменениях.

SEM, или метод моделирования латентных кривых развития (*latent growth curve modeling*), как подход к моделированию лонгитюдных данных предполагает построение структурных моделей латентных изменений и недавно подробно описан О.В. Митиной (2008). Достоинствами этого метода являются его соответствие интуитивным графическим

способам (диаграммам) отображения структурных моделей, возможность операционализации ошибки измерения через построение моделей измерения, в том числе при проверке гипотез об изменениях в латентных переменных (Bollen, Curran, 2006). Тем не менее, этот метод с большим трудом позволяет использовать пропущенные данные и моделировать нелинейные изменения. Как отмечают Р. Плойхарт и Р. Ванденберг (Ployhart, Vandenberg, 2010), использование HLM или SEM подхода в большинстве случаев является вопросом предпочтения конкретного исследователя, и чаще всего результаты использования обоих подходов оказываются практически идентичными. Подробное концептуальное и техническое описание указанных методов можно найти в ставших настольными книгах по анализу иерархических (и в более узком смысле – лонгитюдных) данных С. Рауденбуша и А. Брика (Raudenbush, Bryk, 2002), а также Дж. Зингер и Дж. Уиллета (Singer, Willet, 2003).

Оба метода на сегодняшний день являются де-факто стандартом в лонгитюдных исследованиях за рубежом (отметим, что развитие эти методы получили только в конце XX века), но лишь недавно стали использоваться в отечественных публикациях. Так, двумя уникальными примерами использования иерархического линейного моделирования являются в отечественной литературе исследования С. А. Корнилова, Е. Л. Григоренко и С. Д. Смирнова (2009) и Т. В. Корниловой, С. А. Корнилова и М. А. Чумаковой (2010). В первом лонгитюдном исследовании отслеживалась средняя успеваемость нескольких когорт студентов ( $n=441$ ) на протяжении всего времени их обучения в вузе (9 временных срезов; см. рис. 4). Визуальный анализ данных позволил предположить, что динамика успешности обучения в вузе носит нелинейный характер (см. рис. 4 а), что привело к построению в рамках единой модели отдельных субмоделей изменения успеваемости за четыре временных этапа – с 1 по 3 сессию, с 3 по 5 сессию, с 5 по 8 сессию и с 8 по 9 сессию. При начальной средней успеваемости (при  $t = 1$ ) в 4.63 балла студенты в течение первых трех сессий теряли в успеваемости по .13 балла за сессию, с 3 по 5 сессию изменений в успеваемости не наблюдалось, а с 5 по 8 успеваемость возрастала на .08 балла за сессию. Важным результатом стало выявление на втором уровне модели (см. выше) психологических предикторов исходной успеваемости, темпов ее снижения и возрастания. Так, начальный статус предсказывался уровнем развития измеренных на ранних этапах обучения аналитических способностей студентов, а креативность и практический интеллект выступили факторами, «защищающими» от снижения успеваемости на этапах с 3 по 5 сессию (см. рис. 4 б). Использование HLM методологии позволило провести оценку предикторов динамических характеристик успешности обучения, выйдя за пределы традиционных для отечественной психологии образования моделей исследований.

### Заключение

Сложности использования лонгитюда в психологии связаны со спецификой проверяемых в лонгитюдном исследовании гипотез о развитии, несмотря на синонимичное звучание (гипотезы о развитии), выходящих за пределы одной только психологии развития, но реализуемых в самых различных областях психологии (от клинической до когнитивной). Анализ временной динамики показателей изучаемого процесса в современных исследованиях осуществляется не только в классических схемах лонгитюда, он и дополняет самые различные схемы (от изучения влияния фактора времени на показатели ЗП до моделирования осуществляемых или наблюдаемых влияний – с контролем смешивающихся переменных – в квазиэкспериментальных и собственно экспериментальных исследованиях). При этом к обычным процедурам оценки валидности исследования добавляется ряд специфич-

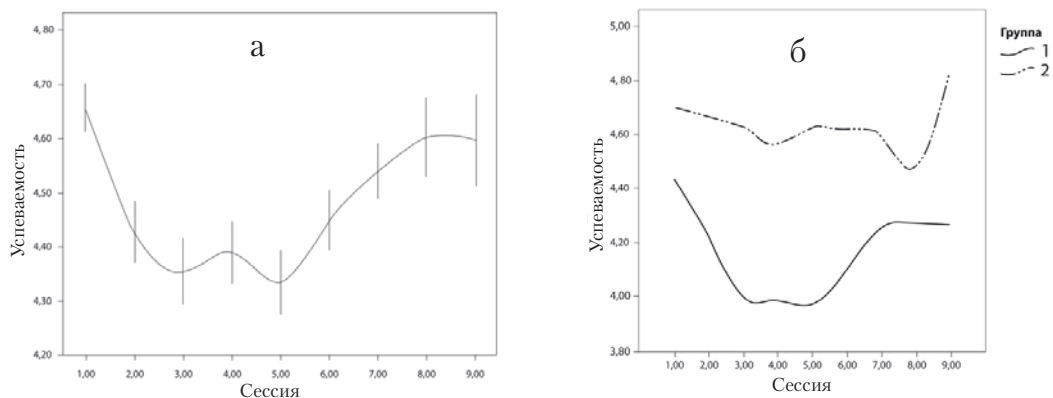


Рис. 4.4 а – средние значения успеваемости студентов с 95-процентным доверительным интервалом; 4 б – график изменения успешности обучения у двух групп студентов (1 – низкие аналитические, творческие, практические способности; 2 – высокий уровень развития всех способностей), основанный на полученных в рамках HLM модели параметрах (и так наз. *fitted values*).

ных для лонгитюда критериев оценки качества исследования.

Таким образом, лонгитюдные исследования – особая группа методов, требующая от исследователя тщательного подхода к планированию схемы исследования и рассмотрения вопросов ее практической реализации, а также серьезной подготовки и скрупулезности в области применения современных статистических методов. Они позволяют выявлять скрытые тенденции, неочевидные кривые развития, которые не всегда можно репрезентировать диаграммами и обычными функциональными графиками. Лонгитюдный метод в современном его звучании предполагает такую спецификацию психологом целей своего исследования, при которой ему совершенно необходимо владеть современными способами многомерной статистики и современными программами, позволяющими моделировать изучаемые зависимости.

### Литература

- Бурменская Г. В. Лонгитюдный метод // Эксперимент и квазиэксперимент в психологии / Под ред. Т. В. Корниловой. СПб.: Питер, 2004. С. 180–192.
- Корнилов С. А., Григоренко Е. Л., Смирнов С. Д. Лонгитюдное исследование академических, творческих и практических способностей как предпосылок успешности обучения // Вопросы психологии. 2009. № 5. С. 54–65.
- Корнилова Т. В. Экспериментальная психология. Теория и методы. М.: Аспект-Пресс, 2003.
- Корнилова Т. В., Корнилов С. А., Чумакова М. А. Лонгитюдное исследование динамики успешности решения студентами аналитических, творческих и практических заданий // Психологическая наука и образование. 2010. № 1. С. 55–68.
- Мутина О. В. Моделирование латентных изменений с помощью структурных уравнений // Экспериментальная психология. 2008. Т. 1. № 1. С. 131–148.
- Baltes P. B., Nesselroade J. R. History and rationale of longitudinal research // Longitudinal research in the study of behavior and development / Eds. J. R. Nesselroade, P. B. Baltes. N.Y.: Academic Press, 1979. P. 61–87.
- Bollen K. A., Curran P. J. Latent curve models: A structural equation perspective. Hoboken, N.J.: John Wiley, 2006.
- Cook N. R., Ware J. H. Design and analysis methods for longitudinal research // Annual Review of Public Health. 1983. № 4. P. 1–24.



- Glenn N. Cohort analysis. Beverly Hills, CA: SAGE, 1977.
- Flynn J. R. What environmental factors affect intelligence: The relevance of IQ gains over time // Current topics in human intelligence. V. 5: The environment / Ed. D. Detterman. Norwood, N.J.: Ablex, 1996. P. 17–29.
- Graham J. W. Missing data analysis: Making it work in the real world // Annual Review of Psychology. 2009. № 60. P. 549–576.
- Graham J. W., Hofer S. M. Multiple imputation in multivariate research // Modeling Longitudinal and Multiple-Group Data: Practical Issues, Applied Approaches, and Specific Examples / Eds. T. D. Little, K. U. Schnabel, J. Baumert. Hillsdale, N.J.: Erlbaum, 2000. P. 201–218.
- Holland P. Statistics and causal inference // Journal of American Statistical Association. 1986. V. 81. № 396. P. 945–960.
- Mason K. O., Mason W. M., Winsborough H., Poole W. K. Some methodological issues in cohort analysis of archival data // American Sociological Review. 1973. № 32. P. 242–258.
- Menard S. Longitudinal Research (2<sup>nd</sup> ed.). Thousand Oaks: SAGE, 2002.
- Menard S. Introduction: Longitudinal research design and analysis // Handbook of longitudinal research / Ed. S. Menard. N.Y.: Elsevier, 2008. P. 3–12.
- Menard S., Elliot D. S., Huizinga D. The dynamics of deviant behavior: A national survey progress report (National Youth Survey Progress Report No. 49). Boulder, CO: Institute of Behavioral Science, 1989.
- Mitchell T. R., James L. R. Building better theory: Time and the specification of when things happen // Academy of Management Review. 2001. № 26. P. 530–547.
- Newcomb M. D., Bentler P. M. Consequences of adolescent drug use: Impact on the lives of young adults. Newbury Park, CA: Sage, 1988.
- Ployhart R. E., Vandenberg R. J. Longitudinal research: The theory, design, and analysis of change // Journal of Management. 2010. V. 36. № 1. P. 94–120.
- Raudenbush S. W. Comparing personal trajectories and drawing causal inferences from longitudinal data // Annual Review of Psychology. 2001. № 52. P. 501–525.
- Raudenbush S. W., Bryk A. S. Hierarchical Linear Models: Applications and Data Analysis Methods (2<sup>nd</sup> ed.). Thousand Oaks, CA: SAGE, 2002.
- Rosenbaum P., Rubin D. The central role of the propensity score in observational studies for causal effects // Biometrika. 1983. № 17. P. 41–55.
- Rubin D. Estimating causal effects of treatments in randomized and non-randomized studies // Journal of Educational Psychology. 1974. № 66. P. 688–701.
- Rubin D. Inference and missing data // Biometrika. 1976. № 63. P. 581–592.
- Ryder N. B. The cohort as a concept in the study of social change // American Sociological Review. 1965. № 32. P. 843–861.
- Singer J. D., Willet J. B. Applied longitudinal data analysis: Modeling change and event occurrence. N.Y.: Oxford University Press, 2003.
- Terman L. M., Oden M. H. Genetic studies of genius: The gifted child grows up. Twenty five years' follow-up of a superior group. Stanford, CA: Stanford University Press, 1947.
- Thomae H. The Bonn Gerontologic Longitudinal Study (BOLSA) // Zeitschrift für Gerontologie, 1993. V. 26. № 3. P. 142–150.
- Twisk J. W. R. Causal inference in longitudinal experimental research // Handbook of longitudinal research / Ed. S. Menard. N. Y.: Elsevier, 2008. P. 279–294.
- Wall W. D., William H. L. Longitudinal studies and the social sciences. London: Heinemann, 1970.
- Willet J. B. Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth // Educational and Psychological Measurement. 1989. № 49. P. 587–602.
- Vandenberg R. J. Toward improving measurement invariance methods and procedures // Organization Research Methods. 2002. № 5. P. 139–158.

## LONGITUDINAL STUDIES: THEORY AND METHODS

*KORNILOV S.A., Lomonosov Moscow State University, Moscow*

The article reviews contemporary approaches to understanding longitudinal designs as a group of methods aimed at testing a specific class of hypotheses (about change, its functional form and dynamic parameters). We provide a description of main longitudinal designs and point out their limitations, along with analyzing major validity threats specific to longitudinal studies: the critical review of methodological and methodical problems and possibilities of overcoming them is completed by the brief description of the main statistical approaches for analyzing longitudinal data.

**Keywords:** longitudinal study, longitudinal design, longitudinal validity, repeated measures hierarchical linear modeling.

### ***Transliteration of the Russian references***

*Burmenskaja G. V.* Longitjudnyj metod // Jeksperiment i kvazijeksperiment v psihologii / Pod red. T. V. Kornilovoj. SPb.: Piter, 2004. S. 180–192.

*Kornilov S. A., Grigorenko E. L., Smirnov S. D.* Longitjudnoe issledovanie akademicheskikh, tvorcheskikh i prakticheskikh sposobnostej kak predposylok uspeshnosti obuchenija // Voprosy psihologii. 2009. № 5. S. 54–65.

*Kornilova T. V.* Jeksperimental'naja psihologija. Teorija i metody. M.: Aspekt-Press, 2003.

*Kornilova T. V., Kornilov S. A., Chumakova M. A.* Longitjudnoe issledovanie dinamiki uspeshnosti reshenija studentami analiticheskikh, tvorcheskikh i prakticheskikh zadanij // Psihologicheskaja nauka i obrazovanie. 2010. № 1. S. 55–68.

*Mitina O. V.* Modelirovanie latentnyh izmenenij s pomow'ju strukturnyh uravnenij // Jeksperimental'naja psihologija. 2008.T. 1. № 1. S. 131–148.