

УДК 37.032:303.7

ПРОБЛЕМА ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ ФАКТОРНОГО АНАЛИЗА В ПЕДАГОГИЧЕСКОМ ИССЛЕДОВАНИИ

Лебедева И.П.

АНО ВПО «Пермский институт экономики и финансов», Пермь, e-mail: kafmos2011@yandex.ru

Рассматриваются способы интерпретации результатов факторного анализа с учетом поставленной задачи исследования и принципа простой структуры, а также ограничения интерпретации конкретного эмпирического материала в связи с возникающими неопределенностями. Обсуждается проблема анализа причинных отношений и возможности латентных факторов в их описании. Приводятся примеры их интерпретации, во-первых, в случае получения релевантных переменных, нагружающих факторы, которые отражают разные стороны реальной ситуации; во-вторых, в случае сравнения факторной структуры для экспериментальной и контрольной групп в формирующем педагогическом эксперименте с целью определения характера причинных отношений на основе латентных факторов. В последнем случае при интерпретации результатов факторного анализа для совокупности экспериментальных и неэкспериментальных данных оценивается степень эвристичности выводов о причинности.

Ключевые слова: факторный анализ, фактор, причинные отношения, неэкспериментальные данные

INTERPERTATION PROBLEM OF THE FACTORIAL ANALYSIS REZULTS IN THE PEDAGOGICAL RESEARCH

Lebedeva I.P.

ANO «Perm institute of economy and finance», Perm, e-mail: kafmos2011@yandex.ru

The article considers the interpretation methods of the factorial analysis results of taking into account of the research tasks and the simple structure principle, and also interpretation restriction of the specific empirical data in connection with the arising uncertainty. The problem of the causal relations analysis and the latent factors possibility in their description is discussed. Examples them interpretation, first, in case of receiving the relevant variables loading factors which reflect the different sides of a real situation are given; secondly, in case of comparing of factor structure for the experimental and control groups in the pedagogical experiment for the purpose of determination of character causal the relations on the basis of latent factors. The uncertainty of conclusions about causality in the course of interpretation of the factorial analysis results of experimental and non- experimental data is estimated.

Keywords: factorial analysis, factor, causal relations, non-experimental data

В педагогическом исследовании в процессе анализа неэкспериментальных данных часто необходимо провести анализ причинных отношений. Для этого используется факторный анализ, представляющий собой многомерный анализ данных, изучающий внутреннюю структуру структуры матриц ковариаций и корреляций [3]. В процессе их преобразования выделяются латентные факторы, детерминирующие значения наблюдаемых переменных и объясняющие взаимосвязи между ними. В итоге выполняется редукция исходных переменных, к которым предъявляются определенные требования для того чтобы получить результаты факторного анализа, адекватно решающие поставленные исследовательские задачи. В содержательном плане различают следующие задачи факторного анализа:

- поиск подтверждения существующей гипотезы, либо формирование новой о существовании латентных факторов (выделенные факторы должны отражать реальные механизмы взаимодействия переменных в изучаемом явлении);
- снижение размерности совокупности переменных;
- классификация объектов.

Решая эти задачи необходимо иметь в виду, что фактор – это искусственный статистический показатель, возникающий в результате специальных преобразований ковариационной или корреляционной матриц. Каждый фактор интерпретируется как причина совместной изменчивости группы переменных, имеющих по этому фактору максимальные по модулю факторные нагрузки. Однако возникает вопрос о реальности этого фактора, поскольку существует бесконечное множество решений, одинаково хорошо объясняющих ковариационную (корреляционную) матрицу.

При выборе из множества эквивалентных решений единственного решения, руководствуясь принципом простой структуры, все равно имеет место ряд неопределенностей:

- достаточно ли велика факторная нагрузка переменной, чтобы причислить ее к данному фактору;
- достаточно ли данного набора переменных с такими факторными нагрузками, чтобы судить о наличии фактора;
- как изменится факторная структура, если учесть корреляцию между факторами [4].

С целью уменьшения этих неопределенностей выполняют проверку устойчивости факторной структуры на разных выборках и вводят маркерную переменную, проявившую устойчивые связи с данным фактором.

Важную роль в интерпретации результатов факторного анализа играет исходный набор переменных, в котором «скрываются» латентные факторы. Формирование такого набора является нетривиальной задачей. Поэтому выделяют две стратегии факторного анализа: эксплораторный и конфирматорный факторный анализ. В любом случае, учитывая факторную структуру, должно быть найдено обобщенное название для группы нагружающихся на фактор переменных. Поскольку переменные могут иметь несколько оснований сходства, то при интерпретации факторов нужно исключить их альтернативные объяснения. При оценке теоретической значимости выделенных факторов имеют в виду, что факторный анализ дает интерпретацию лишь конкретного эмпирического материала и сокращенное описание лишь данного набора параметров. Вычислительные процедуры факторного анализа не предполагают проверку статистических гипотез о параметрах генеральной совокупности, т.е. отсутствует обоснование с точки зрения математической статистики переноса результатов, полученных по выборке, на генеральную совокупность. Обычно она интуитивно каким-то образом понимается исследователем, который допускает при этом определенную эвристичность, снижению которой способствует достаточно большой объем выборки (что и предполагается факторным анализом) и повторение эмпирического исследования на разных выборках.

Однако если результаты факторного анализа интересны, прежде всего, с практической стороны, то они интерпретируются по отношению к исследуемой выборочной совокупности. Практические задачи, не требующие обобщения на другие совокупности данных, часто возникают в различных исследованиях. Тогда полученные факторы имеет смысл анализировать в аспекте реального состояния исследуемого явления, а не их теоретической значимости или научной новизны. Например, при проведении анкетирования часто вопросы (параметры) формируются таким образом, чтобы представить все основные стороны практической ситуации. Факторный анализ будет удачным, если факторы интегрируют релевантные параметры, характеризующие ее конкретную сторону.

Совсем другой случай, когда ставится задача обоснования появления нового фе-

номена в виде выделенного фактора. В качестве предпосылок такого обоснования целесообразно рассматривать специально разработанную теоретическую модель, которой должна соответствовать факторная структура. Заметим, что в некоторых методах факторного анализа (максимального правдоподобия и обобщенного метода наименьших квадратов) предусмотрена проверка гипотезы с помощью критерия хи-квадрат Пирсона о числе факторов [1]. В стратегии конфирматорного факторного анализа предполагается еще и тщательный отбор переменных. Но на начальном этапе исследования в процессе реализации эксплораторного факторного анализа возникает вопрос о «распознавании» феномена, а не о его «рафинировании». И на этом этапе могут возникнуть проблемы:

– противоречие между соответствием принципу простой структуры и теоретической моделью, характеризующей структуру факторов;

– использование методов косоугольного вращения и интерпретация взаимосвязей факторов;

– появление альтернативных факторов (неоднозначность интерпретации факторной структуры).

Первую проблему можно попытаться решить на основе корректировки набора переменных с позиции большего соответствия исходной теоретической модели или совершенствовать саму модель (если она есть). Однако эксплораторный факторный анализ проводят, когда эта модель отсутствует. Ее получение и является задачей факторного анализа, которая решается постепенно путем последовательного улучшения его результатов и проведения повторных эмпирических исследований.

Обратим внимание на то, что встречается ситуация, когда общий фактор выделяется один (факторное решение однозначно), т.е. все переменные оказываются достаточно близкими друг к другу. Как нам представляется, в данном случае нужно начать с постановки задачи и сформулировать ее как интеграцию переменных. При этом важна максимальная информативность результатов факторного анализа.

Открывают возможности интерпретации причинных отношений экспериментальные данные или хотя бы сочетание экспериментальных и неэкспериментальных данных [4, 6]. Проведение факторного анализа с целью определения влияния факторов, сравнения их структуры для разных выборок обучаемых и оценки эффективности действия переменных, обусловивших различия (или, наоборот, идентичность)

структуры факторов в этих выборках, как правило, имеет место в сравнительном эксперименте.

Приведем пример такого эксперимента. Он имел целью оценить меру влияния специального использования систем учебных вопросов как средства обучения на его эффективность. «В эксперименте приняли участие две группы студентов первого курса физического факультета Пермского педуниверситета. Одна из них была контрольная, другая – экспериментальная. Обозначим их как группы *A* и *B* соответственно. В контрольной группе практические занятия по математическому анализу проводились по традиционной для учебного процесса вуза методике, а в экспериментальной группе предусматривалось использование преподавателем систем учебных вопросов разного типа в роли основного стимулирующего и организующего средства учебной деятельности студентов. При выполнении ими организационной функции особое значение имели вопросы методологического характера, ориентированные на использование методов высших уровней иерархии. Ведущее место в стимулировании мышления студентов занимали вопросы, требующие не столько воспроизведения известных фактов, сколько определенных размышлений над сущностью понятий, раскрытия причинно-следственных связей между ними, самостоятельного поиска метода решения серьезной задачи» [2, с. 89].

Применение систем учебных вопросов предусматривалось в различных формах: преподаватель рассказывал о некоторых рациональных схемах и принципах постановки вопросов; демонстрировал их в разных ситуациях; стимулировал студентов к активному оперированию ими. Системы учебных вопросов использовались в следующих типичных ситуациях: обсуждение теоретического материала, решение задач, самостоятельная работа над математическим текстом, контроль знаний.

Эксперимент предусматривал получение оценок результативности проведенного обучения. Для этого были получены показатели, отражающие успешность овладения студентами математикой: качество математической подготовки, уровень интеллекта, математические способности, характеристики некоторых качеств личности. В конце обучения осуществлялось сравнение относительных показателей в каждой группе. Выяснилось, что средние отклонения эффективности (по результатам выполнения контрольных работ и теста интеллекта) и средняя величина абсолютных показателей, измеренных в конце обучения,

выше в экспериментальной группе, чем в контрольной (по критерию Стьюдента, $p < 0,05$). Полученные результаты свидетельствуют о более высокой эффективности обучения математике в экспериментальной группе. Учитывая методику организации эксперимента, можно утверждать, что причиной, оказавшей решающее воздействие на результаты обучения, явилось целенаправленное использование преподавателем систем учебных вопросов в качестве основного стимулирующего и организующего средства мыслительной деятельности студентов. Однако в каждой группе могли действовать и другие факторы, также в определенной мере повлиявшие на конечный результат. Какие это факторы? И какова мера их влияния?

С целью поиска ответов на эти вопросы экспериментальные данные были обработаны методом факторного анализа. Факторизация осуществлялась по методу главных компонент. Суммарная доля воспроизводимой дисперсии в совокупности трех факторами в группе *A* – 57%, в группе *B* – 53%. Структура взаимосвязей в первых двух факторах оказалась в общих чертах сходной для групп *A* и *B*. Первый фактор обозначен «Успешность учебной деятельности», второй фактор интерпретирован как «Дефицит культуры вербально-логического характера», учитывая его отрицательные корреляции с переменными. Третий фактор имеет незначительный вес и недостаточно ясную структуру для интерпретации.

Сравнивая взаимосвязи второго фактора с переменными в каждой группе, можно заметить, что в контрольной группе отрицательных корреляций значительно больше. На основании этого естественно предположить, что дефицит культуры вербально-логического характера в группе *A* более ощутим, включая и умение владеть вопросами. В группе *B* положительные корреляции данного фактора с оценками за выполнение тестов интеллекта в начале вузовского обучения можно предположительно объяснить наличием компенсаторных механизмов, позволяющих с определенной степенью успешности справляться с тестами и без использования сложных умозаключений.

Хотя результаты формирующего эксперимента не позволяют делать вывод об устойчивости факторной структуры (ее проверка не проводилась), и небольшой вес факторов показывает наличие «статистического шума», тем не менее выявление второго фактора в процессе эксплораторного факторного анализа имеет теоретическую и практическую значимость, поскольку объясняет (или хотя бы позволяет выдвинуть

гипотезы) о причинах, которые могут влиять на успешность учебной деятельности студентов. В отличие от неэкспериментальных данных, обнаружение данного фактора в условиях эксперимента свидетельствует о его реальности, что создает возможности функционального контроля второго фактора при проведении различных видов экспериментов (в частности, факторных) и непосредственно в учебном процессе.

Таким образом, интерпретация результатов факторного анализа в случае неэкспериментальных данных позволяет выявить лишь гипотетические причины, детерминирующие взаимосвязи в корреляционной матрице, если ставится задача поиска латентного фактора. Введение маркерной переменной и проверка устойчивости факторной структуры помогает в обосновании предположений о латентных факторах. Их получение можно рассматривать как некую модель исследуемого явления или процесса, для которой возникает вопрос не о ее истинности, а ее теоретической значимости и практической ценности. И даже в этом случае необходимо иметь в виду роль, которую играет исследователь на каждом этапе реализации факторного анализа, поскольку он формирует совокупность переменных и осуществляет выбор конкретных методов факторизации корреляционной матрицы.

Очевидно, что наибольшей субъективностью обладает этап интерпретации выявленных факторов. Кроме обозначенных выше неопределенностей факторного анализа в процессе интерпретации его результатов учитывают, что он предусматривает использование линейных взаимосвязей переменных и факторов. Это ограничивает границы применимости факторного анализа, поскольку построенные модели являются приближением реально существующих сложных нелинейных взаимосвязей для локальных условий. Несмотря на то, что результаты факторного анализа рассматривают как гипотезу о причинности, ее можно считать вполне обоснованной, если исследователь соблюдает требования к его реализации. Кроме того, степень обоснованности такой гипотезы значительно повышается в ситуации, когда:

- неэкспериментальные данные дополняются экспериментальными данными;
- результаты факторного анализа проверяют, т.е. изучают воспроизводимость факторной матрицы в различных условиях (при изменении выборки, повторном измерении переменных и т.д.) и с получением аналогичных данных с помощью других методов многомерной статистики.

В некоторых случаях возникает необходимость использования качественных

переменных, для которых невозможно выполнить требования, предъявляемые к исходным данным. Тогда факторный анализ имеет смысл проводить с поисковыми целями с тем, чтобы в перспективе для качественных данных получить их количественные аналоги. В процессе выполнения факторного анализа только с качественными данными возрастают его неопределенности, которые ставят под сомнение интерпретацию выделенных факторов, поскольку погрешность измерения сказывается на структуре факторов (переменные должны обладать высокой надежностью). Использование частот для номинальных переменных не является способом повышения их надежности, т.к. исходный выборочный ряд значений переменных остается атрибутивным, а формальная их замена на частоты в рамках классического факторного анализа не предусмотрена. Тем не менее подобный разведочный вариант использования факторного анализа целесообразен в случае, если совокупность качественных переменных дополняется количественными показателями. Однако при таком варианте применения факторного анализа сужается круг исследовательских задач, например представляется неэффективным сравнение факторной структуры в разных выборках.

Таким образом, исследовательские задачи и тип исходных данных в значительной мере определяют способы интерпретации результатов факторного анализа, которые можно оценивать либо в аспекте их научной новизны и теоретической значимости, либо с точки зрения практической ценности. Факторное решение является удачным, учитывая требования к научному методу исследования, если все этапы факторного анализа выполнены корректно и однозначность интерпретации факторной структуры достаточно обоснована, что позволяет результаты факторизации рассматривать со всех указанных позиций.

Список литературы

1. Высоков И.Е. Математические методы в психологии: Учебник и практикум для академического бакалавриата. – М.: Изд-во Юрайт, 2014. – 386 с.
2. Лебедева И.П. Математическое моделирование в педагогическом исследовании. Монография. Акад. Акмеол. Наук, Перм. Гос. пед. ун-т. СПб; Пермь: 2003. – 122 с.
3. Лоули Д.Н., Максвелл А.Э. Факторный анализ как статистический метод / Пер. с англ. – М.: Мир, 1967. – 144 с.
4. Наследов А.Д. Математические методы психологического исследования. – 2004.
5. Таганов Д.Н. SPSS: Статистический анализ в маркетинговых исследованиях. – СПб: Питер, 2005. – 192 с.
6. Kolesnikov A.K., Lebedeva I.P. Investigation of causality based on complex use of statistical methods (case study of social research) / Quality&Quantity. 2013. – V. 47, Issue 6. – P. 3043–3050.