

Теоретическая валидизация измерительных шкал

Аннотация

Статья посвящена специфике теоретической валидизации измерительных шкал. Такого рода валидизация предполагает прохождение следующих этапов: подготовка теоретической базы, разработка и корректировка эмпирических индикаторов, проверка конструктивной валидности и ретестовой надежности, описание специфики интерпретации полученных результатов. Основным отличием представленного в статье подхода от классической теории измерения, используемой в социальных науках, является ориентировка на валидную теорию, а не на некую гипотетическую теорию, подлежащую проверке. В контексте используемой (валидной) теории осуществляются все этапы теоретической валидизации — начиная от выбора типа используемой шкалы и заканчивая особенностями интерпретации полученных результатов.

Эмпирический базис статьи — результаты теоретической валидизации комплексной 10-балльной шкалы, предназначенной для измерения степени усилий, прилагаемых студентами в процессе учебы. В рамках разработки данной методики использованы данные одного качественного (изучение феномена обучения студентов в условиях ограниченного времени) и двух смешанных исследований (изучение поведенческих практик студентов, совмещающих работу и учебу, а также анализ влияния организационной структуры малых выпускающих кафедр на особенности обучения студентов).

Ключевые слова: измерение, теоретическая валидизация

Введение

В своих недавних работах [Дембицкий, 2010а; Дембицкий, 2010с; Дембицкий, 2009; Дембицкий, 2008] я затронул широкий спектр вопросов, касающихся проблемы теоретической валидизации. Предложенный мной подход можно резюмировать в следующих положениях:

1. Теоретическая валидность представляет собой степень соответствия теоретического конструкта исследуемому феномену.
2. Теоретическая валидность является комплексной характеристикой, включающей очевидный, содержательный, логический и композиционный компоненты.
3. Теоретическая валидизация в большей мере связана с **достижением** необходимой степени соответствия теоретического конструкта исследуемому феномену, чем с ее проверкой.
4. Завершенный процесс теоретической валидизации связан с обеспечением других видов валидности: конструктивной, внешней, внутренней и прогностической.
5. Теоретическая валидизация должна осуществляться с помощью комплексных исследовательских стратегий.
6. Теоретическая валидизация связана прежде всего с парадигмой исследований смешанного типа (mixed methods research).

В рамках данной статьи демонстрируется роль теоретической валидизации в осуществлении социологического измерения — начиная от критериев построения шкалы и заканчивая спецификой интерпретации полученных данных. В качестве примера используется шкала, предназначенная для измерения усилий, прилагаемых студентами в процессе обучения.

Эта шкала базируется на теоретическом конструкте, сформулированном в рамках исследования, посвященного феномену учебы в условиях ограниченного времени. Конструкт описывает три типа студентов в зависимости от их отношения к учебе: “безразличный”, “декларативно заинтересованный” и “ответственный”. Несмотря на то что теория справедлива в отношении студентов стационара, совмещающих работу и учебу, распространение отдельных ее элементов на всех студентов стационара также справедливо. Соответствующие пояснения представлены ниже.

Первый этап: подготовка теоретической базы

В отличие от общепринятого подхода валидизации измерительных шкал, в рамках теоретической валидизации исследователь должен начинать подготовку измерительного инструмента имея в распоряжении валидный теоретический конструкт, а еще лучше — валидную теорию. В дальнейшем все полученные результаты измерения должны сопоставляться, а в случае необходимости — и модифицироваться в соответствии с этими теоретическими положениями.

При построении шкалы в качестве теоретической базы использованы результаты двух исследований. Первое посвящено поведенческому репертуару студентов или поведенческим стратегиям, второе — упомянутым выше типам студентов.

Остановимся подробнее на содержательных характеристиках этих типов. Каждый из них можно описать с помощью двух поведенческих стратегий — основной и дополнительной (реже встречаются ситуации, когда стратегии имеют равноценное значение). Соответствующие стратегии включают отличительный поведенческий репертуар (см. табл. 1). Описание поведенческих стратегий было получено с помощью исследовательской стратегии согласования концептов [Дембицкий, 2009], позволяющей обеспечить содержательную и логическую валидность теоретического конструкта.

Таблица 1

Содержательные характеристики поведенческих стратегий студентов

Название группы поведенческих практик	Содержание группы
<i>1. Стратегия бездельника</i>	
1.1. “Ничегонеделание”, надежда на удачу	Давать обещания, что в дальнейшем обязательно будешь учиться; использовать моменты, когда преподаватель спешит, и получать списание долгов; быть наглым, идти на риск (ничего не отработать, уходить, не отпросившись, и т.д.); не работать в течение семестра и сдавать долги только на сессии; сдавать несколько раз (брать измором) и т.д.
1.2. “Цель оправдывает средства”	Использовать связи; сказать, что работу забыл дома и попроситься отвечать без нужных материалов; если сдаешь не вовремя, списывать готовые работы и сдавать их; упросить преподавателя, ничего при этом не делая, и т.д.
1.3. Ссылка на внешние причины	Объяснять пробелы и пропуски внешними причинами и “давить” на то, что был вынужден; ссылаться на бытовые проблемы, которые мешают присутствовать; использовать тему здоровья для объяснения пропусков и т.д.
<i>2. Стратегия ловкача</i>	
2.1. Работа с преподавателем	Изучить преподавателя и найти к нему индивидуальный подход; если раньше учился хорошо, использовать свой авторитет; ранжировать преподавателей по требовательности и в зависимости от этого выделять время на подготовку и т.д.
2.2. Использование подходящих ситуаций и моментов	Когда присутствуешь на парах, демонстрировать активность; когда сдаешь долги, перевести обсуждение на материал, который знаешь; готовиться в транспорте и т.д.
2.3. Минимизация усилий	Показывать выполненные задания и отпрашиваться с пар; если есть определенный минимум необходимых заданий, делать только его; если делаешь задания редко, то, по крайней мере, качественно, так как это подкупает; посещать те виды занятий, которые более важны (например, практические, а не лекции) и т.д.
<i>3. Стратегия трудяги</i>	
3.1. Самоорганизация	Планировать, раскладывать по пунктам то, что необходимо осуществить; всячески использовать самоорганизацию и т.д.
3.2. Опора на дружеские отношения в учебной группе	Помогать студентам, которые помогают тебе, распределять подготовку заданий между разными людьми, брать у одногруппников необходимые материалы и т.д.
3.3. Вежливость	Прилежно вести себя на парах; демонстрировать положительное отношение к дисциплине; когда уходишь с пар, всегда согласовывать это с преподавателем и т.д.
3.4. Опора на собственные силы	Всегда стараться вникать в суть материала; если необходимо, оставаться после пар для пересдач; при необходимости тратить все свободное время на подготовку и т.д.

В свою очередь, типы студентов были определены в рамках использования исследовательской стратегии “обоснованной теории” [Дембицкий, 2010а] и могут быть охарактеризованы следующим образом. Если студент относится к “безразличному” типу, он может использовать: а) стратегию бездельника в качестве основной и стратегию ловкача в качестве дополнительной; б) стратегию ловкача в качестве основной и стратегию бездельника в качестве дополнительной (либо в равной степени). Если студент относится к “декларативно заинтересованному” типу, в качестве основной используется стратегия ловкача, а в качестве дополнительной — стратегия трудяги. Наконец, если студент принадлежит к “ответственному” типу, он может использовать: а) стратегии ловкача и трудяги в равной степени; б) стратегию трудяги как основную, а стратегию ловкача как дополнительную.

Несмотря на то что в фокусе исследования находились студенты стационара, совмещающие работу и учебу, полученные данные можно разделить на две части — ту, которая касается исключительно работающих студентов, и ту, которая является общей как для работающих, так и для неработающих. В частности, соотношение типов студентов и соответствующих поведенческих стратегий является справедливым для всех студентов, что обеспечивает высокую внешнюю валидность. Фактически обе группы с точки зрения поведенческого репертуара отличаются лишь в одном — наличием или отсутствием поведенческой группы “Компромиссы на работе”.

Второй этап: разработка и корректировка эмпирических индикаторов в соответствии с используемой теорией

Первоначально на основе этой информации для каждой из указанных стратегий был сформулирован набор индикаторов (см. табл. 2). Все индикаторы оценивались с помощью следующего вопроса: “*Студенты используют различные способы поведения в различных ситуациях, касающихся учебы. Ниже описаны некоторые из этих способов. Скажите, как часто Вы лично используете каждый из них, с применением следующей шкалы: (1) почти никогда не использую — (2) использую довольно редко — (3) порой использую — (4) использую довольно часто — (5) почти всегда использую*”.

Также во время пилотного опроса ($N = 16$) были заданы вопросы относительно пропусков, частоты подготовки к практическим занятиям и успешности сдачи последней сессии. Пилотаж показал, что такой способ определения типа студента является неудовлетворительным.

Во-первых, полученные результаты плохо соотносились с контрольными вопросами относительно частоты посещения занятий, интенсивности подготовки и успешности сдачи сессии. Главной причиной этого стало то, что индикаторы, первоначально предназначенные для измерения конкретной стратегии, могли трактоваться участниками исследования совершенно иначе и, таким образом, не соответствовать той сути, которая первоначально в них вкладывалась.

Во-вторых, результаты, полученные с помощью индикаторов, описывающих одну общую стратегию, имели низкую внутреннюю согласованность. Так, наибольшее значение α Кронбаха, равное 0,44, было зафиксировано для индикаторов, описывающих стратегию бездельника. Для стратегий ловкача и трудяги это значение было равно 0,11 и 0,30 соответственно.

Таблица 2

Набор индикаторов для измерения выраженности поведенческих стратегий

Название стратегии	Содержание индикаторов
Стратегия бездельника	Большое количество долгов, накопленных за семестр, не является для меня серьезной проблемой, ведь их можно сдать и перед самой сессией
	Я сдаю работы, заданные на дом, с отставанием от большей части группы. Это позволяет мне сверить правильность выполнения работы и исправить ошибки
	Свои пропуски я объясняю личными обстоятельствами, которые помешали присутствовать на паре
	Для сдачи многих зачетов/экзаменов мне бывает достаточно ознакомиться с материалами конспекта за день или даже в день сдачи.
	Для того, чтобы получить нужную оценку, я иногда просто вежливо прошу об этом преподавателя
	Когда я объясняю пропуск занятия преподавателю, то использую тему здоровья или семейных обстоятельств
Стратегия ловкача	Когда мне представляется возможность, я демонстрирую учебную активность на паре
	Я стараюсь найти правильный подход к преподавателю, учесть его индивидуальные особенности
	При сдаче долгов я стараюсь перевести обсуждение на материал, который знаю
	Я трачу больше времени для подготовки к предметам более строгих преподавателей
	Я в основном посещаю те виды занятий (лекции либо практические), которые более важны для сдачи предмета на сессии. Остальные занятия посещаю не всегда
	Если преподаватель задает определенный минимум необходимых заданий на семестр, для меня достаточно выполнить только их
Стратегия трудяги	Подготовку домашних заданий мы (я и другие студенты моей группы) распределяем между собой
	Я стараюсь планировать выполнение моих учебных обязанностей
	В учебных целях я прибегаю к сотрудничеству с моими одногруппниками
	Мое поведение на парах является правильным и прилежным
	Я вникаю в суть учебного материала и стараюсь с ним хорошо разобраться
	При необходимости я использую все свободное время для подготовки к учебе

В-третьих, специфика такого измерения не отвечала особенностям имеющейся теории. Дело в том, что согласно теории, разработанной на основе изучения феномена учебы в условиях ограниченного времени, поведенческие стратегии являются своего рода континуумом, начинающимся страте-

гией бездельника, продолжающимся стратегией ловкача и заканчивающимся стратегией трудяги. Отсюда, поведенческий репертуар студента наилучшим образом можно охарактеризовать с помощью отнесения его к двум смежным стратегиям (бездельник и ловкач, ловкач и трудяга), как это описано выше. В процессе исследования с помощью стратегии “обоснованной теории” не было встречено случаев, когда студент использовал все три стратегии в равной степени или же когда наряду с основной стратегией он имел две дополнительных в качестве равноценных. В ситуации измерения с помощью набора индикаторов такие случаи встречались. Гипотетически возможны даже такие ситуации, когда результаты измерения будут принимать максимальное значение для каждой из стратегий, несмотря на то, что такие случаи не фиксируются в эмпирической действительности посредством наблюдения.

На мой взгляд, все это говорит о том, что такой способ измерения является во многом артефактным и, соответственно, непригодным для изучения поведенческих стратегий студентов. Исходя из этого был разработан другой способ измерения на основе одного комплексного вопроса, содержание которого также определялось описанными выше поведенческими стратегиями. С помощью этого вопроса у студентов напрямую спрашивали о том, какую поведенческую стратегию они используют, хотя, по сути, речь шла именно о типе студента:

Давайте предположим, что в самом общем виде студентов можно разделить на три группы, в зависимости от усилий, прилагаемых в процессе учебы:

- А) наименее прилежные студенты, в основном не прилагающие усилий для учебы, надеющиеся на удачное стечение обстоятельств и на то, что все долги можно будет закрыть в конце семестра, а экзамены и зачеты сдать благодаря помощи одногруппников или даже практически ничего не делая;
- В) балансирующие между подготовкой наиболее важных заданий и игнорированием заданий, не имеющих принципиального значения, использующих ранжирование преподавателей по строгости и требовательности с целью соответствующей концентрации усилий и предпочитающих посещать те пары, которые имеют больший вес в рамках дисциплины;
- С) наиболее старательные, относящиеся к учебе максимально ответственно, старающиеся готовить все задания, предусмотренные учебными требованиями, пытающиеся получить максимум знаний и постоянно посещающие пары.

А теперь, используя приведенную ниже шкалу, УКАЖИТЕ МЕСТО, НА КОТОРОЕ ВЫ БЫ ПОСТАВИЛИ СЕБЯ, сравнивая с этими тремя группами. При этом нельзя выбирать “чистую” группу (А, В или С), а необходимо выбрать ОДНО из значений между двумя “чистыми” группами (между А и В или между В и С).

А	A-90%	A-70%	A-50%	A-30%	A-10%	В	B-90%	B-70%	B-50%	B-30%	B-10%	С
	B-10%	B-30%	B-50%	B-70%	B-90%		C-10%	C-30%	C-50%	C-70%	C-90%	

Третий этап: проверка конструктивной валидности и ретестовой надежности

Минимальная проверка конструктивной валидности связана с обеспечением двух ее составляющих — конвергентной и дискриминантной валидности [Campbell, 1959]. Первая основывается на том, что результаты шкалы демонстрируют связь с данными, полученными с помощью методик, предназначенных для измерения концептуально близких социальных феноменов. Вторая же, наоборот, основывается на том, что результаты не связаны с данными методик, предназначенных для изучения концептуально отличных феноменов.

Основное опасение, связанное с использованием представленной выше 10-балльной шкалы, касалось возможности получения социально желательных ответов, то есть вопрос мог оказаться сензитивным и смещать ответы в сторону завышения оценок. Для проверки этого предположения в анкету наряду с другими вопросами была включена шкала лжи из опросника ММРІ.

После удачного пилотного опроса ($N = 24$) шкала была проверена на двух выборках различных вузов ($N_1 = 61$, $N_2 = 62$). Предположительно, в случае адекватного отображения шкалой теоретического конструкта результаты ее применения должны были коррелировать с ответами на вопросы о количестве выступлений на практических занятиях и об успешности сдачи последней сессии, отрицательно коррелировать с ответами на вопрос о количестве пропусков и не иметь корреляции со шкалой лжи. Для соответствующей проверки был использован ранговый коэффициент корреляции Спирмена (см. табл. 3).

Таблица 3

Сила и направление связи шкалы с контрольными переменными

Контрольная переменная	Сила связи		
	Первый вуз ($N = 61$)	Второй вуз ($N = 62$)	Вместе ($N = 123$)
Количество выступлений на практических занятиях	0,74**	0,69**	0,71**
Успешность сдачи последней сессии	0,67**	0,67**	0,67**
Количество пропусков	-0,56**	-0,64**	-0,59**
Шкала лжи	-0,16	0,20 ¹	-0,09 ²

** результаты статистически значимы на уровне 1%.

Исходя из полученных результатов, было принято решение, что данная шкала отображает теоретический конструкт, описывающий поведенческие

¹ Количество человек во втором вузе, опрошенных с помощью Шкалы лжи, равно 17. В дальнейшем данная взаимосвязь проверялась на больших выборках студентов этого вуза, но статистически значимая связь так и не была выявлена ($N = 41$, $r_s = -0,03$).

² $N = 76$.

стратегии студентов, а также, что она слабо подвержена смещениям, связанным с социально желательными ответами.

Вместе с тем приведенные данные касаются проверки шкалы, в первую очередь на уровне самоочета, являющегося во многом субъективным. Поэтому была осуществлена ее дополнительная проверка с помощью экспертов.

В состав последних вошли по два преподавателя и одному студенту на каждой из кафедр. В каждом вузе оценивались все студенты одной учебной группы ($N_1 = 20$, $N_2 = 24$). Преподаватели, отобранные в качестве экспертов, были хорошо знакомы с учебными группами, а студенты-эксперты учились в составе этих групп и работали на этих кафедрах. Оценка производилась с помощью 10-балльной шкалы комплексного вопроса. Кроме этого для каждого из студентов было рассчитано среднее значение оценок на сессиях, полученных ими за все года обучения¹. Для расчета силы связи был использован ранговый коэффициент корреляции Спирмена. Поскольку оценки экспертов в высокой степени согласованы (для первого вуза значение α Кронбаха равно 0,89, для второго — 0,96), дополнительно было рассчитано среднее значение оценок экспертов для каждой кафедры и вычислены соответствующие коэффициенты корреляции (см. табл. 4).

Таблица 4

Сила и направление связи результатов экспертного оценивания со средними значениями оценок, полученных на всех сессиях

Эксперты	Сила связи	
	Первый вуз	Второй вуз
Первый преподаватель	0,83**	0,93**
Второй преподаватель	0,72**	0,91**
Студент	0,80**	0,93**
Среднее значение трех оценок	0,85**	0,95**

** результаты статистически значимы на уровне 1%.

Приведенные данные основываются на более точной информации, чем самоочеты. Во-первых, это связано с согласованностью результатов, полученных в ходе оценивания разными экспертами. Во-вторых, это обеспечивается отсутствием потери информации относительно успешности обучения, поскольку учтены все годы. Следовательно, результаты экспертной проверки полностью подтверждают сделанные ранее выводы.

После проверки конструктивной валидности в трех учебных группах был проведен опрос с целью проверки на ретестовую надежность. Согласованность двух опросов, оцененная с помощью коэффициента корреляции Спирмена, на уровне 0,75 ($N = 37$, вероятность ошибки — 1%).

Четвертый этап: специфика интерпретации полученных результатов

Особого внимания заслуживает специфика интерпретации результатов, полученных при использовании этой шкалы. Дело в том, что содержа-

¹ Были использованы данные об успешности сдачи сессий первых 4-х лет обучения.

тельно шкала основывалась на информации о поведенческих стратегиях студентов, а конечной целью ее использования было получение информации о принадлежности каждого студента к одному из трех типов — “безразличному”, “декларативно заинтересованному” и “ответственному”. Исходя из этого, интерпретация шкалы осуществлялась исключительно на основе теоретических положений, связывающих поведенческие стратегии с типами студентов (см. табл. 5).

Таблица 5

Соотношение значений шкалы с типами студентов

Тип студента	Поведенческие стратегии	Значения шкалы
1. “Безразличный”	1.1. Стратегия бездельника — основная, стратегия ловкача — дополнительная	1, 2
	1.2. Стратегия ловкача — основная, стратегия бездельника — дополнительная или же равноценная	3, 4, 5
2. “Декларативно заинтересованный”	2.1. Стратегия ловкача — основная, стратегия трудяги дополнительная	6, 7
3. “Ответственный”	3.1. Стратегии ловкача и трудяги используются в равной степени	8
	3.2. Стратегия трудяги — основная, стратегия ловкача — дополнительная	9, 10

Следовательно, студенты, выбравшие пять первых градаций шкалы, были отнесены к “безразличному” типу, выбравшие шестую и седьмую градации — к “декларативно заинтересованному”, выбравшие последние три — к “ответственному”.

Для проверки правильности такой интерпретации ее результаты были соотнесены с результатами ответов насчет успешности сдачи сессии (см. табл. 6, 7), предварительно преобразованных в три градации: 1) во время сдачи сессии были тройки; 2) сессия сдана на четыре–пять; 3) сессия сдана на отлично.

Для каждой группы студентов был рассчитан ранговый коэффициент корреляции γ . Для обоих вузов он равен 0,83 и статистически значим на уровне 1%.

Таблица 6

Связь типов студентов с успешностью сдачи сессии (первый вуз)

Результаты сдачи последней сессии	Тип студента			Итого
	“Безразличный”	“Декларативно заинтересованный”	“Ответственный”	
Были тройки	15	6	1	22
Четыре–пять	4	10	2	16
На отлично	1	6	14	21
Итого	20	22	17	59

Таблица 7

Связь типов студентов с успешностью сдачи сессии (второй вуз)

Результаты сдачи последней сессии	Тип студента			Итого
	“Безразличный”	“Декларативно заинтересованный”	“Ответственный”	
Были тройки	21	4	3	28
Четыре–пять	5	11	8	24
На отлично	0	1	9	10
Итого	26	16	20	62

Несмотря на то, что сила и направление связи одинаковы для обоих вузов, ее характер существенно отличается. Однако в рамках данной статьи этот вопрос не рассматривается, поскольку связан не со спецификой самой шкалы, а с особенностями учебного заведения. Этот вопрос рассмотрен в моей публикации [Дембицкий, 2010с]. В целом же данные результаты подтверждают правомерность использования предложенной интерпретации, хотя и указывают на необходимость ее спецификации в отношении высших учебных заведений различных типов, что, в свою очередь, является важным условием достижения композиционной валидности шкалы.

Вместе с тем в контексте группировки пунктов шкалы вполне логичным может выглядеть сомнение касательно приоритетности именно такого варианта группировки. Наиболее очевидными альтернативами являются такие представленные ниже варианты разделения шкалы, которые бы делили ее на приблизительно равные группы.

Первый вариант (3/4/3):

1 группа / “безразличный” тип			2 группа / “декларативно заинтересованный” тип				3 группа / “ответственный” тип		
A-90%	A-70%	A-50%	A-30%	A-10%	B-90%	B-70%	B-50%	B-30%	B-10%
B-10%	B-30%	B-50%	B-70%	B-90%	C-10%	C-30%	C-50%	C-70%	C-90%

Второй вариант (3/3/4):

1 группа / “безразличный” тип			2 группа / “декларативно заинтересованный” тип			3 группа / “ответственный” тип			
A-90%	A-70%	A-50%	A-30%	A-10%	B-90%	B-70%	B-50%	B-30%	B-10%
B-10%	B-30%	B-50%	B-70%	B-90%	C-10%	C-30%	C-50%	C-70%	C-90%

Третий вариант (4/3/3):

1 группа / “безразличный” тип				2 группа / “декларативно заинтересованный” тип			3 группа / “ответственный” тип		
A-90%	A-70%	A-50%	A-30%	A-10%	B-90%	B-70%	B-50%	B-30%	B-10%
B-10%	B-30%	B-50%	B-70%	B-90%	C-10%	C-30%	C-50%	C-70%	C-90%

Если эти варианты группировки являются более удачными, в смысле отображения типа студента, то они должны демонстрировать более высокие, по сравнению с инициальной группировкой (5/2/3), коэффициенты корреляции (γ и Спирмена) с успешностью обучения студентов (см. табл. 8).

Как видно из представленных данных, наиболее эффективным вариантом группировки является вариант 3/3/4, а уже потом 5/2/3. Однако, если сравнить распределения таблиц сопряженности для двух вариантов группировки, то получится противоположная картина (см. табл. 9, 10). В обеих таблицах приводится информация сразу о двух вузах.

Таблица 8

Величина коэффициента γ для различных вариантов группировки пунктов шкалы

Вариант группировки шкалы	Сила связи для первого вуза	Сила связи для второго вуза	Сила связи для обоих вузов*
5/2/3	0,83**	0,83**	0,81**
3/4/3	0,84**	0,80**	0,79**
3/3/4	0,82**	0,89**	0,83**
4/3/3	0,81**	0,78**	0,77**

* массивы объединены;

** результаты статистически значимы на уровне 1%.

Таблица 9

Связь типов студентов с успешностью сдачи сессии (группировка 3/3/4)

Результаты сдачи последней сессии	Тип студента			Итого
	“Безразличный”	“Декларативно заинтересованный”	“Ответственный”	
Были тройки	16	29	5	50
Четыре-пять	1	20	19	40
На отлично	1	3	27	31
Итого	18	52	51	121

Таблица 10

Связь типов студентов с успешностью сдачи сессии (группировка 5/2/3)

Результаты сдачи последней сессии	Тип студента			Итого
	“Безразличный”	“Декларативно заинтересованный”	“Ответственный”	
Были тройки	36	10	4	28
Четыре-пять	9	21	10	24
На отлично	1	7	23	10
Итого	46	38	37	121

Если представить функциональную связь этой зависимости, то логично предположить, что все безразличные студенты имели бы тройки, декларативно заинтересованные учились бы на четыре–пять, а ответственные — на отлично. Исходя из такого критерия, можно оценить, насколько каждый вариант группировки отклоняется от этой идеальной модели. В случае варианта 3/3/4 отклонение наблюдается в 48% случаев (58 из 121), а в случае варианта 5/2/3 — в 34% случаев (41 из 121)¹. В таблице, построенной для варианта 3/3/4, видно, что значительная часть студентов, имеющих тройки, попали в декларативно заинтересованную группу, что не соответствует представленной выше теории. Собственно, именно последний аспект, а не сила установленных взаимосвязей, имеет решающее значение для оценки различных вариантов группировки.

¹ Для вариантов группировки 3/4/3 и 4/3/3 величина отклонения составляет 43% и 42% соответственно.

В целом, более высокие коэффициенты корреляции для варианта 3/3/4 объясняются в данном случае прежде всего особенностями расчета самого показателя γ и, соответственно, несут скорее техническую, чем смысловую нагрузку¹.

Если же для оценки силы связи использовать коэффициент корреляции Спирмена, то группировка 5/2/3 будет более предпочтительной, нежели другие варианты (см. табл. 11).

Таблица 11

Величина корреляции Спирмена для различных вариантов группировки пунктов шкалы

Вариант группировки шкалы	Сила связи для первого вуза	Сила связи для второго вуза	Сила связи для обоих вузов*
5/2/3	0,65**	0,68**	0,66**
3/4/3	0,57**	0,62**	0,58**
3/3/4	0,66**	0,64**	0,64**
4/3/3	0,57**	0,64**	0,59**

* массивы объединены;

** результаты статистически значимы на уровне 1%.

Выводы

Делая выводы, необходимо сказать как о методологических аспектах теоретической валидизации, так и о методических аспектах использования комплексной 10-балльной шкалы, рассмотренной в статье.

В методологическом смысле можно дать две главные рекомендации, касающиеся теоретической валидизации измерительных шкал:

1. Прежде чем переходить к социологическому измерению, необходимо разработать теорию, описывающую конструкты тех социальных феноменов, которые будут объектом измерения. При этом разработанная теория должна быть валидной (то есть соответствовать эмпирической действительности, лежащей в фокусе исследования) уже до измерения. В этом смысле теория является первичной, а измерение — вторичным, производным от нее. Измерительная шкала разрабатывается, модифицируется и оценивается в контексте имеющейся теории. Благодаря этому достигается содержательная и логическая валидность измерительной шкалы, являющиеся базой для достижения конструктивной валидности.

2. После того как шкала прошла проверку на конструктивную валидность, исследователь должен провести измерения в отношении тех социальных групп, для изучения которых пригодна эта шкала. Это необходимо преимущественно для того, чтобы иметь представление об основаниях интерпретации результатов, получаемых при ее использовании. Здесь, опять-таки, невозможно обойтись без валидной теории, вписывающей конструкт, лежащий в основе измерения, в более широкое концептуальное поле. В данном случае уже измерительная шкала (а точнее, ее результаты) используется для обогащения той теории, на основании которой она была построена.

¹ Для детального ознакомления с особенностями расчета коэффициента γ см., напр.: [Хили, 2005: с. 416–420].

В методическом смысле можно рекомендовать следующее¹:

1. Не включать шкалу в анкеты, предназначенные для самозаполнения. Слова Сеймура Садмена о том, что респонденты являются когнитивными скупцами, на все 100% проявляют себя в данном случае. Часто студенты даже не дочитывают инструкцию до конца и отмечают один из описанных в вопросе типов, даже после просьбы внимательно прочитать инструкции вопроса до самого конца. Поэтому желательно, чтобы соответствующая анкета использовалась в face-to-face интервью. Если же проводится групповое интервью (например, с учебной группой), необходимо тщательно объяснять специфику вопроса, а также проверять результаты ответов, чтобы была возможность внести поправки в случае неправильного заполнения.

2. В комплексных образовательных исследованиях, предусматривающих формирование базы данных по каждому студенту, следует опрашивать преподавателей, а не студентов. В данном случае получаемая информация о студентах имеет более точный характер.

3. Следует очень аккуратно относиться к продемонстрированным в данной статье взаимосвязям. Дело в том, что специфика образовательного учреждения может вносить в них существенные коррективы. Последнее имеет важнейшее значение с точки зрения интерпретации полученных результатов. Наилучший способ избежать ошибочных выводов — триангуляция методов (в частности — включенное наблюдение). Из этих соображений описанная выше шкала наиболее пригодна в исследованиях смешанного типа², предусматривающих получение данных из разных источников.

Источники

Дембицкий С. Исследования смешанного типа в социологии / Сергей Дембицкий // Соціальні виміри суспільства. Випуск 2 (13). — 2010b. — С.128–139.

Дембицкий С. “Обоснованная теория”: стратегия сбора и анализа качественных данных при теоретической валидизации / Сергей Дембицкий // Социология: теория, методы, маркетинг. — 2010a. — № 2. — С. 64–83.

Дембицкий С. Применение стратегии согласования концептов в теоретической валидизации (на примере исследования поведенческих практик студентов) / Сергей Дембицкий // Социология: теория, методы, маркетинг. — 2009. — № 4. — С. 99–114.

Дембицкий С. Теоретическая валидизация на различных уровнях социологического исследования / Сергей Дембицкий // Социология: теория, методы, маркетинг. — 2010c. — № 4. — С. 152–178.

Дембицкий С. Теоретическая валидность измерительной процедуры и смещение данных в социологическом исследовании / Сергей Дембицкий // Социология: теория, методы, маркетинг. — 2008. — № 3. — С. 99–118.

Хили Дж. Статистика. Социологические и маркетинговые исследования / Джозеф Хили. — Киев: ООО “ДиаСофтЮП”; СПб.: Питер, 2005. — 638 с.

Campbell D. Convergent and Diskriminant Validation by the Multitrait-multimethod Matrix / Donald Campbell, Donald Fiske // Psychological Bulletin. — 1959. — № 2. — P. 81–105.

1 Приведенные далее рекомендации основываются на опыте, полученном в процессе сбора информации, не отраженном в рамках статьи. Эта информация достаточно привычна с точки зрения проведения прикладных исследований, но весьма важна для тех, кто почувствует необходимость использования шкалы в своем исследовании.

2 Более подробно об исследованиях смешанного типа см., напр.: [Дембицкий, 2010b].