

Психологические науки

УДК: 159.99

DOI:10.24412/2303-9744-2021-4-12-24

ИЗМЕРЕНИЕ ВАЛИДНОСТИ КАК КЛЮЧЕВОЙ ШАГ В РАЗРАБОТКЕ ПСИХОДИАГНОСТИЧЕСКОЙ МЕТОДИКИ

Н.С. Завоеванная

Российский государственный профессионально-педагогический университет
(Екатеринбург, Россия)

Аннотация. В статье рассматривается понятие психометрики в психодиагностике, необходимость психометрического обоснования психодиагностических методик, основные свойства такого обоснования (надежность, валидность, достоверность и репрезентативность), несколько этапов разработки психодиагностической методики и измерение разных видов валидности как один из основных этапов психометрического обоснования методики в процессе ее разработки, модификации или адаптации. Кратко изложена суть содержательной, конструктивной валидности, в том числе конвергентной и дискриминантной, эмпирической валидности по внешнему критерию и прогностическая валидности и способы измерения показателей этих видов валидности психодиагностических методик.

Ключевые слова: *психометрика, дифференциальная психометрика, психометрические свойства надежность, валидность, виды валидности, измерение валидности, достоверность и репрезентативность, психодиагностическая методика, разработка психодиагностической методики*

UDC 159.99

DOI:10.24412/2303-9744-2021-4-12-24

VALIDITY MEASUREMENT AS A KEY STEP IN THE DEVELOPMENT OF A PSYCHODIAGNOSTIC METHODOLOGY

N.S. Zavoevannaya

Russian State Vocation-Pedagogical University (Ekaterinburg, Russia)

Abstract. The article discusses the concept of psychometrics in psychodiagnostics, the need for psychometrics justification of psychodagnostic techniques, the main properties of such justification (reliability, validity, veracity and representativeness), several stages of the development of psychodagnostic techniques and the measurement of different types of validity as one of the main stages of psychometric justification of the methodology in the process of its development, modification or adaptation. The essence of meaningful, constructive validity, including convergent and discriminant, empirical validity by external criterion and prognostic validity and ways of measuring the indicators of these types of validity of psychodagnostic methodology are briefly described.

Keywords: *psychometrics, differential psychometrics, psychometric properties reliability, validity, types of validity, measurement of validity, veracity and representativeness, psychodagnostic methodology, development of psychodagnostic methodology*

Психодиагностические методики очень разные и цели их создания разные: одни создаются в развлекательно-познавательных целях и направлены на развлечение или актуализацию познавательной активности испытуемого; другие – разрабатываются в профессиональных целях, т.е. направлены на получение объективной и достоверной информации о состоянии тех или иных сфер личности испытуемого, либо уровня развития его способностей, знаний, умений и навыков, а также отдельных качеств и чертах характера и др. Одним словом, профессиональные психометрические

психодиагностические методики¹ направлены на получение необходимой специалисту информации для принятия того или иного решения, постановки диагноза

Естественно, такой инструмент должен быть психометрически обоснован, как отмечает А.Г. Шмелев [12]. Иными словами он должен отвечать требованиям достоверности, надежности, валидности, т.е. давать правдивые, объективные, стабильные соответствующие конкретной предметной области, ответы на вопросы психодиагноста. Эти требования озвучены целым рядом авторов – и зарубежных, и отечественных специалистов в области психодиагностики. Например, А. Аннастази, С. Урбина [1], Ф.Л. Шмидт, Дж. Хантер [13], А.Г. Шмелев [12], Н.А. Батурин и другие [10]. Они могут быть достигнуты или нет. Оценка достижения этих требований – это психометрическая проверка психодиагностической методики – один из ключевых этапов разработки профессионального инструмента.

При этом *психометрика* или *психометрия*² может быть определена в самом общем виде как область психологии, связанная с теорией и практикой измерения психических явлений [3]. Или может быть определена, например, как совокупность теоретико-математических моделей и процедурно-методических правил организации сбора и обработки эмпирических данных, позволяющих выразить психические свойства и параметры психических процессов в числовой или квазичисловой (ранговой, категориальной) форме. [6]

Очевидно, что приведенные выше определения блики друг к другу. Однако и первое определение, приведенное Л.Ф. Бурлачуком и С.М. Морозовым[3], и второе, - приведенное в словаре под редакцией Л.А.Карпенко, А.В. Петровского, М. Г. Ярошевского [6] очень широкие. Они отражают взгляды исследователей на общую психометрику. Тогда как психодиагносту обычно нужно более прикладное определение. И, как правило, применительно к психодиагностике под психометрикой исследователи понимают, так называемую, *дифференциальную*

¹ Здесь и далее имеется в виду именно психометрические методики.

² Понятие было введено еще в первой половине XVIII столетия Х. Вольфом.

психометрику. Либо же сводят воедино общую и дифференциальную психометрику, акцентируя внимание на психометрических свойствах психодиагностической методики.

В отечественной психодиагностике, например, *психометрика*, в частности *дифференциальная*, часто понимается как:

- область психометрии, определяющая и обосновывающая требования к измерению индивидуально-психологических различий в психологической диагностике (Л.Ф. Бурлачук, С.М. Морозов) [3];
- науку о дифференциально-психологических измерениях, выступающую по отношению к психодиагностике и дифференциальной психологии в статусе технолого-методической дисциплины: она обосновывает требования, которым должны удовлетворять измерительные психодиагностические методы, обосновывает процедуру их разработки и применения (В.В. Столин, А.А. Бодалев) [7];
- область психометрики, которая выражает в количественной форме индивидуальные психологические различия между людьми, в ней разработаны основные критерии качества психологического измерения — такие свойства психологических тестов, как надежность, валидность, репрезентативность [6];
- раздел психологии, изучающий теоретические и методологические проблемы психологических измерений; занимается разработкой математических моделей психологического измерения (модель Терстоуна, модель многомерного шкалирования, модель латентных черт), регламентирует экспериментальную проверку по критериям валидности, надежности и прочих выявляемых психологических свойств [9];
- область психологии, которая изучает теоретические и методологические проблемы измерений во всех других областях психологии разрабатывает математические модели для методов психологического измерения (например, модель Терстоуна, модель многомерного шкалирования, модель латентных черт,

факторный анализ); определяет формальные требования для экспериментальной проверки психометрических свойств (валидности, надежности и др.) разнообразных методов психологического, в том числе психодиагностического, измерения [2].

Таким образом, *психометрика* - это направление научно-прикладных исследований в области психодиагностики, имеющие своей целью разработку надежных, валидных и достоверных и репрезентативных новых психодиагностических инструментов и проверку уже имеющихся, т.е. повышение объективности психодиагностических обследований и исследований в целом. И если кратко излагать, то:

- 1) *надежность* – это помехоустойчивость результатов измерения (отвечает на вопрос психодиагноста «Насколько можно быть уверенным, что я получу такие же результаты при следующем измерении?»);
- 2) *валидность* – это соответствие результатов измерения заявленной области исследования (отвечает на вопрос психодиагноста «Насколько можно быть уверенным, что я измерил именно то, что хотел?»).
- 3) *достоверность* – это соответствие реальности полученных результатов (отвечает на вопрос психодиагноста «Насколько можно быть уверенным, что результаты, которые я получил, соответствуют действительному положению дел?»);
- 4) *репрезентативность* – это соответствие выборки стандартизации контингенту, на который нацелена методика (отвечает на вопрос психодиагноста «Насколько можно быть уверенным, что нормы, полученные мной на выборке стандартизации, могут быть применены при диагностике других людей этого же возраста?»).

Одним словом, очевидно, что, как мы уже не раз утверждали на страницах этой статьи и других работах [4; 11], психодиагностический инструмент должен соответствовать критериям качества, выработанными психологами-экспертами. И понятно, что одним из главных критериев

качества является соответствие психометрическим требованиям к методике. И ключевой шаг здесь – измерение валидности.

Объяснение этому простое: можно и нужно убедиться в устойчивости результатов, их достоверности и репрезентативности, но какой в этом смысл, если нет уверенности в предмете диагностики, т.е. психодиагност сомневается в том, что именно он измерил? – Очевидно, смысла в этом нет. И психометрические свойства всегда должны «играть в квартете», если можно так выразиться. И «ведущим музыкантом» в этом «квартете» всегда является валидность.

При этом из всего множества определений валидности [1; 2; 3; 6; 7; 10] нам близко следующее. «Валидность – это мера соответствия результатов тестирования заявленной цели тестирования, в частности тому свойству (или свойствам), которые измеряются» [13, с. 430]. Не смотря на то, в данном определении А.Г. Шмелев говорит о валидности применительно к тестированию – как методу психодиагностики, – его вполне можно отнести и к другим психодиагностическим инструментам.

Таким образом, она раскрывает то, насколько психодиагност будет близок к цели – изучить предмет обследования или исследования, применяя ту или иную методику. И именно поэтому валидность – главное психометрическое свойство. И каждый разработчик психометрического инструмента должен об этом помнить и измерять это свойство у своего инструмента. Эта мысль и проста и лежит на поверхности.

Однако, к сожалению, по долгу профессиональной деятельности [5], раз за разом приходится встречать работы, посвященные разработке и/или адаптации того или иного психодиагностического инструмента без учета показателей отдельных видов валидности [1; 3; 6; 12]:

- содержательной или экспертной;
- конструктивной – конвергентной и дискриминантной;
- эмпирической валидности по внешнему критерию;
- прогностической.

Конечно, видов валидности больше, чем перечислено выше [1; 3; 6; 12]. Тем не менее, если не уделять должное внимание показателям перечисленных видов валидности – всех или отдельных, – разрабатываемой

методики, то она получает научно-популярный характер и может применяться только в научно-исследовательских или познавательно-развлекательных целях. Другими словами, она выполняет скорее «декоративную» функцию, чем функцию изучения предмета психодиагностики, в совокупности с другими методиками и методами, с целью вынести обоснованное психологическое заключение, либо получить действительные результаты исследования группы людей.

Тем самым, дискредитируется и обесценивается вся работа по разработке нового психодиагностического инструмента или адаптации, либо модификации, уже существующего. А, это, необходимо отметить, не такая простая работа. В самом общем виде она может быть *представлена следующими шагами:*

- 1) определение цели, т.е. определение, на что направлена будет методика, какую теоретическую основу под собой будет иметь и какой она будет иметь вид (например, опросник или тест);
- 2) разработка модели методики – определение объема и характера стимульного материала (например, сколько будет шкал и субшкал, и сколько пунктов будет содержать шкала³, какие это будут пункты – прямые или обратные, в каком соотношении);
- 3) разработка стимульного материала⁴ – формирование содержания методики (например, формулирование пунктов опросника или их перевод, в случае если речь идет об адаптации методики, их экспертный отсев);
- 4) апробация методики – пилотное исследование на выборке соответствующей контингенту испытуемых методики (например, мужчины и женщины в возрасте от 16-ти до 55-ти лет, психические здоровые);

³ К этому пункту разработки методики, в свою очередь, также предъявляются определенные требования (смотрите информацию о них, например, в «Российском стандарте тестирования персонала» [10].)

⁴ Под стимульным материалом здесь мы понимаем содержание любой методики.

- 5) редакция модели и стимульного материала методики на основе результатов ее апробации;
- 6) повторная апробация методики;
- 7) психометрическое обоснование методики (измерение надежности, разных видов валидности, достоверности и репрезентативности).

Получается, что все «ломается» уже на последнем этапе. В тот момент, когда огромная часть работы уже проделана. Причин у этого печального факта может быть несколько, но наш взгляды есть *4 причины* (на самом деле, их точно намного больше и в каждом конкретном случае *наиболее острых*):

- пресловутое отставание отечественной психологической науки и практики от западной;
- развивающаяся культура оказания психологической помощи (психодиагностика – это неотъемлемый этап оказания психологической помощи);
- разработка методики только в научно-исследовательских целях;
- трудоемкость измерения отдельных видов валидности.

В рамках данной статьи, нас более всего интересует предполагаемая четвертая причина. Итак, действительно, например, *содержательная валидность*, т.е. ее вид, подразумевающий пригодность стимульного материала методики для цели диагностики, измеряется практическим путем достаточно трудоемко: 1) необходимо найти, как минимум трех экспертов в интересующей психодиагностика области; 2) передать им материалы методики и организовать процедуру независимой оценки; 3) посчитать согласованность оценок экспертов; 4) на основе согласованной оценки экспертами стимульного материала методики, сделать вывод об ее содержательной валидности. И в это причина того, что, как правило, содержательная валидность подтверждается теоретически – на основе теории, которая ложится в основу методики.

Близка к содержательной валидности и, так называемая, *конструктивная валидность* – соответствие содержания методики тому психологическому конструкту, на который она нацелена. Однако если разработчик говорит уже о том или ином психологическом конструкте (например, личностной тревожности), а не просто о содержании стимульного материала, значит он может говорить об ее *эмпирическом измерении*:

- во-первых, проверяя *коррелируют ли значимо шкалы с общим результатом*, т.е «работают» ли шкалы на один и тот же результат, доказывая или опровергая наличие единого психологического конструкта;
- во-вторых, проверяя *коррелируют ли значимо шкалы со шкалами других валидных методик*, направленных на измерение и изучение того же конструкта, либо же НЕ коррелируют значимо со шкалами, направленными на измерение противоположного конструкта, т.е. ожидаемо либо давая связи с со схожими шкалами, свидетельствуя о *конвергентной валидности*, либо НЕ давая связи с противоположными шкалами, свидетельствуя о *дискриминантной валидности*.

Таким образом, главная трудность в последнем случае – подобрать уже валидизированные, психометрически обоснованные методики. Одни словом, разработчик может не так сложно получить эмпирическое доказательство валидности своего инструмента.

Если же говорить об *эмпирической валидности по внешнему критерию* – практической пригодности методики, выражающемся в показателе корреляции данных методики с определенным критерием и/или критериальным поведением, то основная трудность здесь заключается в выборе этого критерия и его источника. Очень сложно подобрать максимально объективный источник информации, поэтому всегда лучше использовать несколько источников информации (документов, данных других методов и методик и т.д.).

Близка к этой сложности и парадоксальная ситуация с *прогностической валидностью*, т.е. потенциалом методики предсказывать поведение испытуемого в будущем, выражающаяся в показателе связи

данных и методики и будущего критериального поведения (его наличия или отсутствия в будущем). Иными словами, разработчик, вынужден ждать событий, которые подтвердят или опровергнут данные, полученные при пилотном исследовании. Например, он вынужден ждать наступит ли депрессия у тех, испытуемых, у которых он изучал склонность к ней при разработке нового тест-опросника склонности к депрессии. Однако это делает инструмент максимально профессиональным и дорогим, поэтому вполне можно заплатить такую цену.

Таким образом, есть определенные трудности на пути получения сведений о валидности методики. Однако ничего сверхестественного в преодолении этого пути от разработчика не требует. Более того, этот путь может быть преодолен без каких-либо специальных средств обработки данных. Ярким примером тому для нас служит большая работа Сергея Рэмовича Панталева над «Методикой исследования самоотношения», проведенная им в 1988-1999 гг [8]. Очевидно, что в те годы речи не шло о компьютерных программах, подобных SSPS for Windows или Rstudio. Тем не менее, она была проделана... В целях получения профессиональной психометрической методики.

В целом, получение профессиональной психометрической методики каждым из разработчиков – это из явных способов преодоления кризиса в психодиагностике. Одним словом, психологическая наука и практика только обогатятся, если каждый разработчик той или иной психодиагностической методики пойдет этим путем. А, требования реальной практики рано или поздно приведут к этому стремлению каждого действительно заинтересованного и увлеченного работой разработчика. Вопрос о том, сколько именно на это потребуется времени, остается открытым...

Литература:

1. Анастаси А., Урбина С. Психологическое тестирование. 7-ое издание. СПб.: Питер, 2009. 688 с.
2. Большой психологический словарь: словарь / под ред. Б. Г. Мещерякова, В. П. Зинченко. СПб.: Прайм-ЕВРОЗНАК; Москва: ОЛМА-пресс, 2003. 666 с.

3. Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психодиагностике. 2-е изд., перераб. и доп. СПб: Питер, 2004. 520 с.
4. Завоеванная Н.С. Соблюдение авторского права по отношению к психодиагностическому и иному инструментарию (о проекте многокритериальной онлайн-экспертизы психодиагностических методик) // Наука. Мысль. 2016. № 8-1. С. 104-109.
5. Завоеванная Н.С. Указатель «Методика-автор»: Комплекс ПД-триада (Л.И. Переслени, Е.М. Мастюкова, Л.Ф. Чупров, М.С. Певзнер, 1990) // Наука. Мысль. 2014. № 7. С. 52-57.
6. Краткий психологический словарь / под общей редакцией А. В. Петровского, М. Г. Ярошевского ; составитель Л. А. Карпенко. 2-е изд., расш., испр. и доп. Ростов-на-Дону : Феникс, 1998. 505 с.
7. Общая психодиагностика. / А.А. Бодалев, В.В. Столин. СПб.: «Речь», 2004. 440 с.
8. Пантилеев. С.Р. Методика исследования самооотношения. – М.: «СМЫСЛ», 1993. 32 с.
9. Психометрика // Академик. Словари и энциклопедии на Академике. URL: <https://dic.academic.ru/dic.nsf/psihologic/1493> (дата обращения: 30.01.2022).
10. Российский стандарт тестирования персонала / Федерация оценки персонала // Организационная психология. 2015. Т. 2, № 2. С. 67-138
11. Чупров Л.Ф., Белобрыкина О.А., Хохлов Н.А., Завоеванная Н.С. О Критериях качества психологической экспертизы (по материалам четырех дискуссионных сетевых площадок) или психологический "Балет дроздования" в шести действиях с эпилогом. // Вестник по педагогике и психологии Южной Сибири. 2019. № 1. С. 13-41.
12. Шмелев А.Г. Практическая тестология / А.Г. Шмелев. М: Маска, 2013. 688 с.
13. Schmidt F.L., Hunter J.E. The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. 1998Psychological Bulletin 124(2): 262-274 p. DOI:10.1037/0033-2909.124.2.262

References:

1. *Anastazi A., Urbina S.* Psihologicheskoe testirovanie. 7-oe izdanie. SPb.: Piter, 2009. 688 p.
2. Bol'shoj psihologicheskij slovar': slovar' / pod red. B. G. Meshherjakova, V. P. Zinchenko. SPb.: Prajm-EVROZNAK; Moskva: OLMA-press, 2003. 666 p.
3. *Burlachuk L.F., Morozov S.M.* Slovar'-spravochnik po psihodiagnostike. 2-e izd., pererab. i dop. SPb: Piter, 2004. 520 p.
4. *Zavoemannaja N.S.* Sobljudenie avtorskogo prava po otnosheniju k psihodiagnosticheskomu i inomu instrumentariju (o proekte mnogokriterial'noj onlajn-jekspertizy psihodiagnosticheskikh metodik) // Nauka. Mysl'. 2016. № 8-1. pp. 104-109.
5. *Zavoemannaja N.S.* Ukazatel' «Metodika-avtor»: Kompleks PD-triada (L.I. Peresleni, E.M. Mastjukova, L.F. Chuprov, M.S. Pevzner, 1990) // Nauka. Mysl'. 2014. № 7. pp. 52-57.
6. Kratkij psihologicheskij slovar' / pod obshhej redakciej A. V. Petrovskogo, M. G. Jaroshevskogo ; sostavitel' L. A. Karpenko. 2-e izd., rassh., ispr. i dop. Rostov-na-Donu : Feniks, 1998. 505 p.
7. Obshhaja psihodiagnostika. / A.A. Bodalev, V.V. Stolin. SPb.: «Rech'», 2004. 440 p.
8. *Pantileev. S.R.* Metodika issledovanija samootnoshenija. M.: «SMYSL», 1993. 32 p.
9. Psihometrika // Akademik. Slovarei i jenciklopedii na Aka-demike. URL: <https://dic.academic.ru/dic.nsf/psihologic/1493> (data obrashhenija: 30.01.2022).
10. Rossijskij standart testirovanija personala / Federacija ocenki personala // Organizacionnaja psihologija. 2015. T. 2, № 2. pp. 67-138
11. *Chuprov L.F., Belobrykina O.A., Hohlov N.A., Zavoemannaja N.S.* O Kriterijah kachestva psihologicheskoy jekspertizy (po materialam chetyreh diskussionnyh setevyh ploshhadok) ili psihologicheskij "Balet drozdovanija" v shesti dejstvijah s jepilogom. // Vestnik po pedagogike i psihologii Juzhnoj Sibiri. 2019. № 1. pp. 13-41.
12. *Shmelev A.G.* Prakticheskaja testologija / A.G. Shmelev. M: Maska, 2013. 688 p.
13. *Schmidt F.L., Hunter J.E.* The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of

Research Findings. 1998 Psychological Bulletin 124(2): 262-274 p.
DOI:10.1037/0033-2909.124.2.262

— ● —

Сведения об авторе

Наталья Сергеевна **Завоеванная**, педагог-психолог, Российский
государственный профессионально-педагогический университет
(Екатеринбург, Россия); e-mail: natasha_z_ya@mail.ru

— ● —