How to state general qualitative facts in Psychology?

Stéphane Vautier

Quality & Quantity manuscript No.
(will be inserted by the editor)

How to State General Qualitative Facts in Psychology?

Stéphane Vautier

Received: date / Accepted: date

Abstract In what form should qualitative psychological statements have to be in order to be general and falsifiable? I show how basic qualitative measurement systems can be defined that allow for the detection and testing of general relational facts of the if-then form in a well-defined range of potential qualitative, static or dynamic, multivariate observations. The formal framework permits the pointing out of general research problems falling under the scope of empirical investigation.

Keywords Qualitative research · Measurement · Falsifiability

1 Introduction

Empirical scientists search for empirical statements that are general in a sense that, supposedly, they can be fully specified. Moreover, empirical scientists praise falsifiable statements, because recognizing error gives an opportunity to revise existing knowledge [1,6]. Nevertheless, it does not seem straightforward to find examples of general and falsifiable qualitative statements in psychology. In the present paper, I investigate how general qualitative facts could be formulated in such a way that it is made clear how to check and improve their empirical validity, reserving random variables for interpretative purposes.

The first section criticizes the paradigm which consists of the routine statistical operationalization of intuitive claims. Usual quantitative intuitions expressed in so-called general hypotheses are not falsifiable without the help of operationalization. Operationalization of a 'general hypothesis' involves a double process: quantitative encoding, and statistical restatement [8]. The second section lays bare the formal structure of the domain of observables, and focuses on the if-then rule as a special form of a general qualitative relational factual statement. In the third section practical

S. Vautier
Université de Toulouse-Le Mirail, Pavillon de la Recherche, 5 allées Antonio Machado, 31058 Toulouse cedex 9
E-mail: vautier@univ-tlse2.fr

aspects of if-then rules, like their detection and what to do with falsifying cases, are outlined.

2 Problems with Common Trends in Psychological Research

2.1 Statistical Operationalization of Intuitive Claims

A problematic trend in psychological research consists of dealing with statements which do not fulfill the falsifiability criterion. For example, consider the following statement:

> $\mathscr{H}$: Smokers trying to stop smoking feel more anxious than smokers not trying.

Asking in what sense could such a statement be valid, one realizes that there is no obvious answer; "to feel more anxious than" has no natural associated validity criterion. Nevertheless, the statement refers purportedly to a general empirical fact, since its applies to any smoker.

The statistical operationalization of that statement $\mathscr{H}$ would cause the investigator to estimate parameters of a random 'anxiety' sampling variable Y conditionally to an independent variable X assessing the status of a smoker. The experimental approach would randomly assign smokers to experimental ($X = 1$) vs. control ($X = 0$) conditions, whereas the observational approach would merely observe the statut of each smoker on the independent variable X. In both cases, the null hypothesis of equal means would be tested, and likely rejected—statement $\mathscr{A}$:

$$\mu(Y|X = 1) > \mu(Y|X = 0),\ p < .05,$$

where the parameters to be estimated are defined with respect to normal density functions of the same variance in an ad hoc probability space; moreover, an effect size could be computed, indicating the magnitude of the mean effect—statement $\mathscr{B}$, say: $\hat{d} = .3$.

Are $\mathscr{A}$ and $\mathscr{B}$ general statements? Statements $\mathscr{A}$ and $\mathscr{B}$ are not intended to characterize particular smokers at any time. It seems that statistical statements are specific to random variables defined with respect to a specific probability space, which cannot be falsified.

Moreover, the ANOVA approach to a statement like $\mathscr{H}$ has been considered a methodological paradox, as

> In the physical sciences, the usual results of an improvement in experimental design, instrumentation, or numerical mass of data, is to increase the difficulty of the "observational hurdle" which the physical theory of interest must successfully surmount; whereas, in psychology and some of the allied behavior sciences, the usual effect of such improvement in experimental precision is to provide an easier hurdle for the theory to surmount ([3], p. 103).

The paradox would vanish if the focus statement was represented by the null hypothesis instead of its alternative, that is, the numerical values of $\mu(Y|X)$ were explicitly hypothesized. Unfortunately, numerical values cannot be derived from quantitative intuitions like $\mathscr{H}$.

2.2 Measurement Scales and Implicit Nominal Approximations

Another problematic trend in psychological research, which is closely related to the former, consists of the preference for quantitative-like formulations, despite a lack of evidence that there is something quantitative to be measured [2,4,5]. Numerical encoding of 'anxiety' is paradigmatic. A set of questionnaire items is selected, and self-ratings are then assigned a numerical value, yielding a composite test score, which is thought of as a more or less reliable and valid measurement of a 'relative position' on the 'anxiety' continuum—test scores allegedly vary on an interval scale. However, test scores do not even possess the property of representing completely ordered sets of responses.

One example is sufficient to highlight this point. Consider a very simple test Y composed of two dichotomic items $a,\ b$. Each item refers to a psychological process or concept allowing the ordering of the observed values: $0_i \leq_i 1_i$, $i = a,\ b$. The set $\mathrm{M}(Y)$ of the possible response vectors is

$$\mathrm{M}(Y) = \{(0_a,\ 0_b),\ (0_a,\ 1_b),\ (1_a,\ 0_b),\ (1_a,\ 1_b)\}.$$

Most pairs of response vectors in $\mathrm{M}(Y)$ can be ordered by the relationship $\leq_{ab}$, which is defined as follows:

$$(x_a \leq_a y_a \text{ and } x_b \leq_b y_b) \Leftrightarrow (x_a,\ x_b) \leq_{ab} (y_a,\ y_b), \tag{1}$$

where $\{x,\ y\} = \{0,\ 1\}$. However, $\leq_{ab}$ is a partial order because it cannot order the pair $\{(0_a, 1_b),\ (1_a, 0_b)\}$.

Now, consider the usual test scores associated with the test Y. They are elements in the range of the aggregation function f, which assigns to any response vector of $\mathrm{M}(Y)$ the algebraic sum of its numerical values. According to Stevens [7], in order to use test scores as measurements, their numerical properties should reflect the empirical properties of the objects they represent, that is, the vectors y_i of $\mathrm{M}(Y)$, $i = 1,\ \ldots,\ 4$. But $f(y_i)$ does not reflect validly the empirical properties of the vectors in $\mathrm{M}(Y)$. In particular, it is contradictory to state that the vectors $y_2 = (0,\ 1)$ and $y_3 = (1,\ 0)$ are distinct on the one hand, which means that they do not belong to the same nominal class, and to state, on the other hand, that $f[(0,\ 1)] = 1$ and $f[(1,\ 0)] = 1$ which means that $(0,\ 1)$ and $(1,\ 0)$ belong to the same nominal class.

From the representational point of view, test scores are replete with *nominal approximations*. Whether nominal approximations can be justified is not the point. The point is that, to the best of my knowledge, test developers do not justify nominal approximations. Hence, usual numerical psychological scales have no clear epistemological foundations. In the next section, I develop a qualitative measurement perspective which does not require the kind of nominal approximations implied by the "quantitative" interpretation of composite test scores, and which provides a clear meaning of the word "general".

3 Restrictive If-Then Rules in Multivariate Qualitative Measurement Systems

3.1 The Multivariate Qualitative Measurement System

A multivariate qualitative measurement system is composed of a vector

$$Z = (Z_1,\ \ldots,\ Z_k)$$

of $k \geq 1$ empirical variables, where at least one variable is qualitative, i.e., of the nominal or ordinal type. Notice that the term "variable" does not refer to a probability space, but is used to denote empirical values referring to an empirical attribute. If some variables of Z are quantitative, there is no loss of generality to treat them formally as qualitative—ordinal—variables. Each variable Z_i, $i \in \{1,\ \ldots,\ k\}$, is characterized by its set of possible values $\mathrm{M}(Z_i)$. Thus, values of observables range in the set $\mathrm{M}(Z)$, which is the Cartesian product of the sets $\mathrm{M}(Z_i)$:

$$\mathrm{M}(Z) = \mathrm{M}(Z_1) \times \cdots \times \mathrm{M}(Z_k). \tag{2}$$

A multivariate qualitative measurement system may also comprise *change variables*. A change variable refers to a couple of (static) variables $(Z_{i1},\ Z_{i2})$ associated with two time points 1 and 2, and its values are vectors in $\mathrm{M}(Z_{i1}) \times \mathrm{M}(Z_{i2})$. Such a change variable can be labeled a 2-time variable. More generally, one can introduce x-time variables in Z, where x denotes the number of time points.

The cardinality of $\mathrm{M}(Z)$ is given by

$$\mathrm{card}[\mathrm{M}(Z)] = \prod_{i=1}^{k} m_i, \tag{3}$$

where m_i is the cardinality of $\mathrm{M}(Z_i)$. In practice, $\mathrm{card}[\mathrm{M}(Z)]$ may be a huge number. I will elaborate on this point later.

The complete definition of the measurement system requires the definition of its domain, viz., the population U of individuals who are supposed to possess a value in $\mathrm{M}(Z)$ at time $t \in T$, where T denotes an appropriate time interval.[1] Coming back to the smokers' example, should John, who smokes no more than 10 cigarettes a year, be included in the population of 'smokers'? The inclusion problem can be restated by defining the population U as the population of animals; then, being or not being a 'smoker' pertains to the specification of Z, by introducing in Z the variable, say, Z_1, which assigns a value to each individual with respect to what is meant by 'smoker'—Z_1 may be a Cartesian product of variables allowing the convenient specification of a 'smoker'.

3.2 Restrictive If-Then Rules

Given a well-defined measurement system $[U,\ T,\ \mathrm{M}(Z)]$, what kind of general facts can be formulated? I will focus on relational facts, that is, on statements specifying

[1] When measurements refer to x time points, t may be defined as a convenient vector in T^x.

associations between an independent, multivariate variable X, and a dependent, possibly multivariate variable Y. Thus, the vector Z decomposes as the couple $Z = (X,\ Y)$, and the codomain $\mathrm{M}(Z)$ decomposes as the Cartesian product $\mathrm{M}(Z) = \mathrm{M}(X) \times \mathrm{M}(Y)$.

Various kinds of relational facts can be defined with respect to the measurement system $[U,\ T,\ \mathrm{M}(X) \times \mathrm{M}(Y)]$. I will focus on factual statements having the form of restrictive if-then rules. Restrictive if-then rules in $[U,\ T,\ \mathrm{M}(X) \times \mathrm{M}(Y)]$ refer to observations performed on any individual at any time in an appropriate time interval $t \in T$. They express formally as

$$\begin{aligned} \exists\ \alpha(X) \subset \mathrm{M}(X),\ \beta(Y) \subset \mathrm{M}(Y),& \\ \forall\ (u,\ t) \in U \times T,& \\ X[(u,\ t)] \in \alpha(X) \Rightarrow Y[(u,\ t)] \in \beta(Y),& \end{aligned} \tag{4}$$

which reads

> at least a strict subset $\alpha(X)$ of values on X and a strict subset $\alpha(Y)$ of values on Y exist, such that for any pair of an individual u observed at time t in the appropriate time interval T, falling in the class associated with the subset $\alpha(X)$ implies falling in the class associated with the subset $\alpha(Y)$.[2]

Such rules can be noted $\mathscr{R}_{\alpha}^{\beta}$. They are general in the sense that they are supposed to be valid for any $(u,\ t) \in U \times T$. They are restrictive in the sense that they apply to strict subsets of values $\alpha(X)$ and $\beta(Y)$. The conditions $\alpha(X)$ allow the specification of a subpopulation of individuals and possible experimental conditions affecting their responses on Y.

For example, consider the statement "each time a smoker tries to stop smoking, she or he get anxious." Let $Z = (X_1,\ X_2,\ Y)$, where X_1 takes the value 1 when u is a smoker, 0 otherwise, X_2 takes the value 1 when "u's trying to stop smoking" is valid, 0 otherwise, and Y takes the value 1 when u endorses the item questionnaire "I feel anxious", and 0 otherwise. $X = (X_1,\ X_2)$. Thus, expressing the statement as the restrictive if-then rule $\mathscr{R}_{(1,\ 1)}^{1}$ yields

$$\mathscr{R}_{(1,\ 1)}^{1} :\ \forall\ (u,\ t) \in U \times T,\ X(u,\ t) = (1,\ 1) \Rightarrow Y(u,\ t) = 1, \tag{5}$$

which makes it a falsifiable statement, provided (i) the decision rules used to assign a value in $\mathrm{M}(Z)$ to any individual are objective—objective rules are rules which yield rater-invariant measurements in $\mathrm{M}(Z)$—, and (ii) there is no logical impossibility to observe a counterexample.

3.3 Falsification of a Restrictive If-Then Rule

To falsify $\mathscr{R}_{(1,\ 1)}^{1}$, it is sufficient to find one counterexample, that is, one occurrence of the form

$$\neg\mathscr{R}_{(1,\ 1)}^{1} :\ \exists\ u \in U,\ \exists\ t \in T,\ \ X(u,\ t) = (1,\ 1) \text{ and } Y(u,\ t) = 0. \tag{6}$$

[2] In the sequel, the notation $X[(u,\ t)]$ will be simplified by the notation $X(u,\ t)$.

Whether formal falsification should entail the withdrawal of the statement in the set of "interesting" statements will be discussed later.

The preceding development gives the form of a restrictive if-then empirical rule on $[U,\ T,\ \mathrm{M}(X) \times \mathrm{M}(Y)]$. Suppose that an investigator is able to state the restrictive if-then empirical rule $\mathscr{R}_i^\beta$:

$$\begin{aligned} \exists\, i \in \{1,\ \ldots,\ \mathrm{card}[\mathrm{M}(X)]\}, \\ \exists\ \beta(Y) \subset \mathrm{M}(Y), \\ \forall\ (u,\ t) \in U \times T, \\ X(u,\ t) = x_i \Rightarrow Y(u,\ t) \in \beta(Y). \end{aligned} \tag{7}$$

$\mathscr{R}_i^\beta$ is falsified by

$$\begin{aligned} \exists\ (u,\ t) \in U \times T, \\ X(u,\ t) = x_i \text{ and } Y(u,\ t) \in \mathrm{M}(Y) \backslash \beta(Y). \end{aligned} \tag{8}$$

However, considering a given data sample, $\mathscr{R}_i^\beta$ cannot be falsified if no $(u,\ t)$ can be observed such that $X(u,\ t) = x_i$. Whether some $(u,\ t)$ can be observed such that $X(u,\ t) = x_i$ is an empirical issue. Reporting the evidenced validity of a given restrictive if-then rule requires two empirical data: (i) the number n of relevant cases that have been observed, and (ii) the number e of exceptions that have been discovered. Thus, a given if-then rule $\mathscr{R}_i^\beta$, which is a theoretical, and general statement about observables, can be associated with its empirical counterpart $\mathscr{R}_i^\beta(n,\ e)$.

Obviously, being able to report general facts is a prerequisite to being able to explain them. The next section focuses on data analysis oriented toward the finding of potentially interesting findings in an observed contingency table associated with the measurement codomain.

4 Practicing Restrictive If-Then Rules

Given a measurement system $[U,\ T,\ \mathrm{M}(X) \times \mathrm{M}(Y)]$, the contingency table associated with a relevant data sample is useful to detect two kinds of interesting regularities, viz., sufficient conditions, and necessary conditions.

4.1 Detection of Potential Sufficient Conditions

Let the rows and columns of the observed contingency table be associated with the values x_i on X, and the values y_j on Y, respectively, where $i = 1,\ \ldots,\ \mathrm{card}[\mathrm{M}(X)]$ and $j = 1,\ \ldots,\ \mathrm{card}[\mathrm{M}(Y)]$. Let n, $n_{i\bullet}$, $n_{\bullet j}$, and n_{ij} denote the sample size, the frequency of the row i, the frequency of the column j, and the frequency of the cell ij, respectively.

Potential restrictive if-then rules correspond to interesting rows in the contingency table. An interesting row i is defined as a row such that

$$\begin{aligned} n_{i\bullet} > 0, \\ \exists\ j \in \{1,\ \ldots,\ \mathrm{card}[\mathrm{M}(Y)],\ n_{ij} \approx 0, \end{aligned} \tag{9}$$

where the sign $\approx$ means that the cell frequency n_{ij} can be viewed as no quantity or negligible quantity of exceptions falsifying $\mathscr{R}_i^j$. The sufficient condition associated with $Y(u,\ t) = y_j$ is $X(u,\ t) = x_i$.

The screening of the entire contingency table associated with the data may provide a series of findings of the form $\mathscr{R}_i^j(n_{i\bullet},\ n_{ij})$, where $n_{ij} \approx 0$, that are eligible for introduction into the list of basic nomothetic statements relevant to $\mathrm{M}(X) \times \mathrm{M}(Y)$. Whether such general empirical statements are of substantive import is a theoretical issue.

To illustrate, consider the following observed contingency table associated with the bivariate independent variable $X = (X_1,\ X_2)$ and the univariate dependent variable Y.

	0	1
00	1	4
01	1	0
10	6	0
11	2	0

(10)

A potential restrictive if-then rule is $\mathscr{R}^0_{\neg(0,\ 0)}$; in other words, a potential sufficient condition associated with $Y(u,\ t) = 1$ is negatively formulated as $X(u,\ t) \neq (0,\ 0)$.

4.2 Detection of Potential Necessary Conditions

The data also suggests a necessary condition associated with $Y(u,\ t) = 1$, that is, $X_2(u,\ t) = 0$. This necessary condition can be stated as the following restrictive if-then rule:

$$X_2(u,\ t) \neq 0 \Rightarrow Y(u,\ t) \neq 1, \tag{11}$$

which is equivalent to

$$Y(u,\ t) = 1 \Rightarrow X_2(u,\ t) = 0. \tag{12}$$

More generally, necessary conditions associated with a given subset $\delta(Y)$ of states on Y are states $\gamma(X)$ on X such that

$$Y(u,\ t) \in \delta(Y) \Rightarrow X(u,\ t) \in \gamma(X). \tag{13}$$

The systematic detection of necessary conditions on X associated with a given state on the dependent variable is out of the scope of the present paper.

5 Discussion

Psychological research that rests on so-called general hypotheses operationalized through ad hoc probability spaces is unlikely to discover general relational empirical facts, if the word "general" means that what is stated is valid for any $(u,\ t) \in U \times T$ [8]. Furthermore, a basic requirement of scientific inquiry is making clear what can be observed, which is hardly achieved with pseudo-quantitative measurements. Starting from a domain of observables based on a finite set of independent and dependent possibly multivariate qualitative attributes, I have shown that empirical statements having the form of restrictive if-then rules are general statements. Actually, restrictive if-then rules define special relations $(\mathrm{M}(X),\ \mathrm{M}(Y),\ G)$, where G is a graph such that there is at least one possible row with at least one impossible cell. The words "possible" and "impossible" refer to observations $(u,\ t)$ drawn in $U \times T$.

The range within which restrictive if-then rules can be stated and falsified is explicit. An investigator interested in finding potential restrictive if-then rules in the empirical world of interest only has to search for "gaps" in actualized rows of the observed contingency table associated with $\mathrm{M}(X) \times \mathrm{M}(Y)$—actualized "gaps" suggest impossible cells in possible rows. Also, restrictive if-then rules can serve to define necessary conditions associated to a given state $\beta(Y)$ on the dependent variable Y.

Sufficient or necessary conditions are *theoretical statements* because they are supposed to be valid for any case $(u,\ t)$ which can be observed (measured). Contrary to what happens in physics, psychology stumbles over identity of people, which could be understood as a sufficient condition for making generality impossible. It is noteworthy that the formulation displayed in Equation 4 allows for different levels of generality. Levels of generality are partially ordered because they refer to conjunctions of restrictions on the codomain of X. For example, adding an identity variable in X allows for the selection of a specific person in U. Thus, the generality of sufficient or necessary conditions that are restricted to a given person lies on T. Then, generality is possible as soon as T is not a point.

Restrictive if-then rules state a *necessity* principle: if some conditions $\alpha(X)$ are fulfilled on X, it is necessary that other conditions $\beta(Y)$ be fulfilled on Y too. Thus, one can wonder whether it is reasonable to expect valid restrictive if-then rules in psychology. There are some reasons to expect that psychological if-then rules may be formulated with a high degree of corroboration: robust effects do exist in psychology.

Although one counterexample suffices to falsify a restrictive if-then rule, it would be unwise to automatically deny the interest of that rule. As the result of a measurement process, falling in a given observational class may be considered a fallible observation. Thus, exceptions may result from measurement error. However, suspecting measurement error should yield the investigator to repeat his or her observation rather than to interpret automatically the exception as the effect of randomness. Counterexamples can be analyzed according to their epistemic value and their practical impact in decision settings.

From the epistemic viewpoint, counterexamples may be trivial or non-trivial. Trivial counterexamples are those that could be easily eliminated if additional, excluding conditions were taken into account, yielding a more detailed characterization of $(u,\ t)$ at the level of the independent variables—and a new if-then rule—, or those that could be eliminated if problems in measurement turned out to explain the exceptions. Non-trivial counterexamples are those cases which cannot be eliminated within the current state of knowledge. Can unknown excluding independent variables be discovered? Can the tuning of the dependent variable be improved in order to eliminate the exceptions?

From the practical viewpoint, a rule of type $\mathscr{R}_{\alpha}^{\beta}$ (see Equation 9) allows one to predict $Y(u,\ t) \in \beta(Y)$ given $X(u,\ t) \in \alpha(X)$. Let n be the number of observed cases falling in the class $\alpha(X)$, and e the number of relevant counterexamples. The predictive interest of the rule $\mathscr{R}_{\alpha}^{\beta}(n,\ e)$ is threatened by the ratio e/n, which gives the proportion of counterexamples to the rule. This ratio can be used as an estimate of a probability to be wrong when predicting $Y(u,\ t) \in \beta(Y)$ given $X(u,\ t) \in \alpha(X)$, if the probability space underlying such an interpretation is specified appropriately.

Sufficient or necessary conditions associated with a restrictive if-then rule may be interpreted in terms of causality. However, causal interpretations require available general facts. An important task of the scientific endeavor in psychology is to find corroborated general facts, and facts of the form of restrictive if-then rules are potential candidates. If a given measurement system $[U,\ T,\ \mathrm{M}(Z)]$ turns out to be unproductive, that is, if everything is possible in $\mathrm{M}(Z)$, the scientific challenge becomes: find another form of general facts that can be stated in $[U,\ T,\ \mathrm{M}(Z)]$, or find a new observational domain $\mathrm{M}(Z')$ within which restrictive if-then rules can be found.

Acknowledgements I am grateful to Didier Dubois and Henri Prade for their helpful comments on a previous version of the present paper.

References

1. G. Bachelard, La formation de l'esprit scientifique [The formation of the scientific mind]. Vrin (12th ed., original work published 1938), Paris (1983)
2. C. Essex and W. E. Smythe, Between numbers and notions: A critique of psychological measurement, Theory & Psychology, 9, 739-767 (1999)
3. P. H. Meehl, Theory-testing in psychology and physics: A methodological paradox, Philosophy of Science, 34, 103-115 (1967)
4. J. Michell, Normal science, pathological science and psychometrics, Theory & Psychology, 10, 639-667 (2000)
5. J. Michell, Is psychometrics pathological science?, Measurement: Interdisciplinarity Research and Perspectives, 6, 7-24 (2008)
6. K. R. Popper, The logic of scientific discovery. Basic Books, Oxford England (1959)
7. S. S. Stevens, On the theory of scales of measurement, Science, 103, 667-680 (1946)
8. S. Vautier, The operationalization of general hypotheses versus the discovery of empirical laws in Psychology, Philosophia Scientiae (in press)