Eur Arch Psychiatr Neurol Sci (1984) 234 : 86–88

European Archives of Psychiatry and Neurological Sciences

Statistical Manoeuvres in the Dark

Remarks on "An Operationalized Multisymptomatic Model of Neuroses (OMMON): Toward a Reintegration of Diagnosis and Treatment in Behaviour Therapy" by I. Hand and W. Zaworka

Matthias Burisch and Detlef Rhenius

Universität Hamburg, Fachbereich Psychologie, von-Melle-Park 5, D-2000 Hamburg, Federal Republic of Germany

Summary. The paper by Hand and Zaworka (1982) is examined in detail. It is concluded that most of their findings are not substantiated by the data presented.

1. Introduction

In a recent paper in this journal, Hand and Zaworka (1982) have suggested a four-dimensional model of neuroses, backed by data from 216 clinical cases, mostly of the obsessional-compulsive variety. Essentially these data consist of self-ratings made on four different instruments.

The article is not particularly clear in exposition. After condensation, however, four specific claims are made by the authors:

(1) Neurotic patients with only one isolated complaint are comparatively rare; most neuroses are complex.

(2) There is a "main route" or typical sequence of stages from no through to high psychopathology. Obsessional-compulsive symptoms are the earliest to appear, followed by depression, social introversion, and phobias, in this order. The final stage is reached when a patient exhibits symptoms in all four categories.

(3) Using data from a single assessment only, specific predictions as to the further development of individual patients can be derived from the model.

(4) The model permits causal inferences regarding the symptomatology, i.e. is capable of identifying the "causes" of a particular symptom configuration.

We do not wish to take issue with the first of these claims but maintain that the remainder is totally unfounded, impressive data handling notwithstanding.

2. Analysis

2.1 Multisymptomatic Model of Neuroses

Using self-rating questionnaires and an admittedly arbitrary cut off point, Hand and Zaworka point out that 58% of their 216 clinical cases present complaints in at least two areas according to their definition. Should this finding come as a surprise to behaviour therapists, given the vast amount of evidence showing that measures of psychopathology tend to correlate, we have no problem with it. The technical quality and clinical utility of the instruments used is a separate matter and need not concern us here. What prompted our comments are the additional claims made by the authors.

Offprint requests to: M. Burisch at the above address

2.2 Main Individual Route of Psychopathology

Hand and Zaworka's Fig. 1 contains a graphic presentation of the 16 possible symptom configurations (called "stages") together with their empirical frequencies in the sample. It can immediately be seen that stages 1, 5, 11, 15, and 16 are the most frequent. These stages correspond to "no pathology" (1), "obsession-compulsion" (5), "obsession-compulsion plus depression" (11), "obsession-compulsion, depression, and social inhibition" (15), and finally, "all four" (16).

Although Hand and Zaworka are not very explicit regarding the purpose of Fig. 1 and the text contains plenty of disclaimers, their description of the figure and the sheer use of terms such as "route", "stage", and "development" implies a dynamism which is simply not present in cross-sectional data. They seem to assume, e.g. that the most likely last or next stage for a patient in stage 11 is either 5 or 15, simply because these are the most frequently occupied "neighbouring" categories. In other words, they estimate transition probabilities where there is no evidence of transition whatsoever. Closer analysis reveals that there is nothing in the data to suggest that a person in 11 was in 5 before and will be in 15 next. It would have been easy to check the appropriateness of the model using the longitudinal data the authors mention as being available. Pending confirmation of this sort, the arrows in the model are purely speculative.

Some simple considerations show that even the alleged "main route" is not indicative of any conspicuous trend. In Hand and Zaworka's system, a patient gets a + sign on any one of their four dimensions if his or her score exceeds the mean score of the respective clinical reference group. As can be read from their Fig. 1, 157 patients (72.7%) satisfied this condition for obsession-compulsion, 97 (44.9%) for depression, 87 (40.3%) for social inhibition and 58 (26.9%) for phobia. Entering the foregoing proportions as unconditional probabilities and assuming stochastic independence for + signs on all four dimensions, it is easy to calculate the expected frequencies for the 16 "stages" in Table 1.

Obviously, there are several stages where the observed frequencies depart from the predicted ones, most notably stages 1 and 16. This reflects the fact that the assumption of stochastic

Table 1. Predicted and observed frequencies for the 16 stages of OMMON

Stage	*n* predicted	*n* observed
1	14.20	32
2	5.21	8
3	9.58	8
4	11.57	4
5	37.79	39
6	3.52	3
7	4.25	0
8	13.87	7
9	7.81	4
10	25.48	15
11	30.80	30
12	2.87	0
13	9.35	7
14	11.31	9
15	20.77	26
16	7.63	24
	216.01	216

independence is overly simplistic; Hand and Zaworka mention scale intercorrelations between 0.20 and 0.47. What this simple exercise demonstrates, however, is that even a random model perfectly reproduces the "main route" 1-5-11-15-16 (in fact, expected and observed frequencies correlate at rho = 0.79). In other words: The heavy traffic along the "main route" signifies little more than the composition of the sample (mostly obsessive-compulsives), the cut off points, and the scale intercorrelations.

2.3 Individual Developments over Time

We have already pointed out that the apparent movement in Hand and Zaworka's Fig. 1 is spurious; the alleged snapshot can as well depict a still life. The authors' VACOPA procedure goes one step further in the direction of data animation. Formerly, an individual's profile had been coded into a configuration of + and − signs using a single cut off per dimension. With the advent of VACOPA, no less than three cut off points per dimension are employed, which yield three configurations ("stages") for each profile. Depending on the extremeness of the profile, the three codes can be identical or not. The former case would be read by the authors as "indicating stable monosymptomatic ... or stable multisymptomatic... pathology", the latter as resembling "a continuous sequence of steps through the model" or a move "from no/very low to higher pathology... or... from pre-existing to even higher pathology" (p. 374). The sobering truth, however, is that a static profile of four questionnaire scores does not in itself possess "stability" or "movement", even if it is viewed against three different backgrounds. The authors concede that their procedure does not "allow safe predictions regarding the direction of possible developments (routes)", but this appears to be the only conceivable purpose it was created for.

The absurdity of the rationale behind VACOPA can best be recognized by way of an analogy. Assume that mean weight in a sample of obese patients is 100 kg (SD = 15 kg), mean body temperature in a group of "feverish" patients is 39.5° C (SD = 1°), mean diastolic blood pressure in a sample of hypertonics is 110 mm Hg (SD = 10 mm Hg) and that mean blood sugar in a group of diabetics is 130 mg% (SD = 5 mg%). We are now in a position to create a General Integrated Medical Model of Internal Complaints (GIMMIC) which will immediately be amalgamated with VACOPA. Assume further that Patient A, who has influenza, has the following profile: 75 kg; 40.1° C; 130 mm Hg; 100 mg%, while Patient B, an overweight diabetic, has a profile of 95 kg; 37.5° C; 112 mm Hg; 140 mg%. Table 2 shows what GIMMIC and VACOPA would do to these miserable persons' data. Using Hand and Zaworka's logic would make us label Patient A as a case of "stable pathology", whereas some imminent changes would be predicted for Patient B's condition!

2.4 Causal Symptom Interactions

The concept of causality, especially if applied to neuroses, is an intricate one. Nonetheless, the authors claim to find sufficient information in their data "to derive causal interactions of symptoms".

It must be said in all fairness that their treatment of the issue is probably the most confusing part of the paper. VACOPA is again said to be responsible for generating hypotheses, but there is no evidence (or necessity) for this use. Moreover, the heuristic value of the endeavour seems quite limited: Almost anything goes! The only possibilities declared as impossible are (a) unidirectional cause-effect relationships between two equally strong symptoms, (b) reciprocal causal interdependence between a stronger and a weaker symptom, and (c) a weaker symptom causally influencing a stronger one. These exclusions, however, are only by assumption; it is conceded that subsequent tests can prove (a) wrong; finally, it is

Table 2. Simulation of two hypothetical patients' courses through stages of GIMMIC (with varying cut off point procedure)

Patient A						Patient B					
Cut off point	O	F	H	D	Stage	Cut off point	O	F	H	D	Stage
$\bar{X} + \frac{1}{2}$ SD	−	+	+	−	9	$\bar{X} + \frac{1}{2}$ SD	−	−	−	+	5
$\bar{X}$	−	+	+	−	9	$\bar{X}$	−	−	+	+	11
$\bar{X} - \frac{1}{2}$ SD	−	+	+	−	9	$\bar{X} - \frac{1}{2}$ SD	+	−	+	+	14
Individual course 9.9.9						Individual course 5.11.14					

Symptoms: O = Obesity
F = Fever
H = Hypertonia
D = Diabetes

possible under any circumstances that two symptoms have nothing to do with each other and there is no guidance as to when this hypothesis should be preferred. We would even venture to question the plausibility of assumption (c): Why should not, e.g. the realization that he begins to develop phobic fears worsen a patient's moderately severe depression or social inhibition?

Apparently it has also escaped the authors' attention that no such mysterious device as VACOPA was necessary to arrive at their conclusions. Simply converting their raw scores to a common scale would have done the same job more efficiently.

Our main objection, however, is that it is just not good enough to "derive hypotheses", particularly if there are data at hand to test them. Any first term student will be happy to contribute a hunch if asked to do so. What is needed is evidence to support hypotheses.

3. Conclusion

In summary, then, we eagerly await the results section of publications Hand and Zaworka mention as forthcoming. Until then, most of their promises remain to be fulfilled. In their own words: "On methodological grounds it seems impossible to draw conclusions from one single assessment about the development and the interdependence of variables over time." It *is* impossible. Period.

References

Hand I, Zaworka W (1982) An Operationalized Multisymptomatic Model of Neuroses (OMMON): Toward a reintegration of diagnosis and treatment in behaviour therapy. Arch Psychiatr Nervenkr 232:359–379

Received October 28, 1983