Brunswik-Symmetry and Successfully Predicting Human Behavior

Werner W. Wittmann

Universität Freiburg
Federal Republic of Germany

August 1988

Paper presented at the XXIVth International Congress of Psychology, Sydney, Australia, August 28 - September 2, 1988, as part of the Symposium on The Predictability of Human Behavior.

Tuesday, August 30, 8:30 AM - 11:30 AM

Running Head: Brunswik-Symmetry

Brunswik-Symmetry and Successfully Predicting Human Behavior [a]

Werner W. Wittmann
Universität Freiburg
Federal Republic of Germany

Abstract

Successes in predicting human behavior are reported that lead to unusually high coefficients as compared to conventional wisdom in personality research. Using f.e. Eysenck's factors extroversion (E) and neuroticism (N) measured via the Freiburg Personality Inventory (FPI) a correlational pattern is shown which clearly fits the Campbell & Fiske criteria of convergent and discriminant validity. The pattern is as follows: FPIE with FPIN -.25, the criteria for E(CE) and N(CN) correlate CE with FPIE .70, CN with FPIN .82, CE with FPIN -.26, and CN with FPIE -.15. Usual threats to validity of scientific conclusions as outliers, capitalization on chance etc. are controlled and cannot be taken as an explanation. It is demonstrated that successes as these are related to principles of symmetry derived from Brunswik's representative design, Tucker's lens model equation, a multivariate reliability and validity theory, depending heavily on partitioning Cattell's data box and multiple act criteria.

[a] Attending the congress and presenting this paper was possible through the help of the Deutsche Forschungsgemeinschaft DFG (Az 477/3/88).

Brunswik-Symmetry and Successfully Predicting Human Behavior

Werner W. Wittmann
Universität Freiburg
Federal Republic of Germany

1. The significance of prediction in science

Why must we predict human behavior? Answer: To detect the regularities of human life and its governing laws! How well must we predict? Answer: Better than chance! How much better than chance must we be able to predict? Answer: It depends. If you want to survive in the social-darwinistic battles between the different branches of psychological theories in terms of research funds, esteem and reinforcement from your scientific peers, your predictions should be better than those of your competitors. Why do we make so much fuss about prediction? We all know from epistemology that prediction is only a necessary but not sufficient condition for explanation. Successful explanation is the ultimate goal of science. Successful prediction does not automatically imply successful explanation. Psychology knows that quite well, the early successes in prediction, in a period coined as "dust-bowl--empiricism", did not pay off in successful explanation. But is there successful explanation without successful prediction? Yes, there is a lot of it, if we agree to define successful as convincing, i.e. explanations which are believed by people. Yet in science at the heart of successful explanation lies repeatability. Explanation without testable prediction is religion, art, literature, is cargo-cult-science (Feynman 1985) at its best. Psychology as religion, as art, as literature is very, very interesting but not science.

Mischel's message was clearly understood. Prediction from personality as regards trait or psychodynamic psychology was obviously poor, "... with the possible exception of intelligence, highly generalized behavioral consistencies have not been demonstrated, and the concept of personality traits as broad dispositions is thus untenable" (Mischel 1968, p. 146).

Fig. 1 depicts the evaluation standards of empirical research in psychology, and the two dangers related to minimal and maximal predictability.

Fig. 1 about here

There was great hope in the camp of behaviorism especially the social cognitive learning theory brand of Stanford (Bandura 1977) that situational variables will do a better job of prediction and explanation. Learning theory had been very successful in terms of asterisks of significance (but see Meehl 1978). In the same year when Mischel's book appeared, another seminial paper was published namely: "Multiple regression as a general data-analytic system" (Cohen 1968). Cohen pointed to the unnecessary division of psychology's research tradition in an experimental and a correlational branch and how the general linear model synthesizes both. It was now easily possible to transform the asterisks of significance into correlation coefficients as effect sizes. The scientific community rapidly recognized that situational variables did not account for greater parts of behavioral variance than dispositional ones.

Darley and Batson (1973) in their brilliant and ingenious experiment "From Jerusalem to Jericho" found only hurry, a situational variable, to be a significant predictor of helping behavior. The effect size was r_{pb} =.42. The stage was prepared for interactionism. Interestingly the majority of situationists converted to interactionism most prominently Walter Mischel (Mischel 1977). "Traits are alive and well" Epstein (1977) marks the beginning of serious counterattack from trait psychology. In a continuous series of papers Seymour Epstein (1979, 1980, 1983a, 1983b, 1984, 1986, Epstein & O'Brien 1985) brought some important lessons to our mind which seemed to have been forgotten by many researchers, namely the virtues of reliability and the different variants of aggregation to improve on reliability and validity. The situation/disposition/interaction-controversy changed into the consistency controversy. Mischel no longer denied the virtues of aggregation but still challenged the cross-situational consistency (Mischel & Peake 1982), i.e. the lack of predictability from traits across situations. The controversy and the different remedies are comprehensively described in Manfred Amelang and Norman Endlers papers, so I can shorten with only wondering about the logical inconsistency of interactionists denying traits! With what shall those situational variables interact? Bandura (1986) is more consistent with the old attacks on trait psychology. He does not deny that averaging across situations, occasions and forms of behavior may enhance predictability. "Aggregation inflates correlations but yields indeterminate or empty predictions" (Bandura 1986, p. 10). With statements like these Bandura echoes the earlier criticism on "dust-bowl-empiricism" which may lead to good predictions but no explanation and by implication no comprehensive theory.

Combining better predictability with better explanation is the crucial point in psychological theory. The remedy I have proposed is therefore best described as theory-directed aggregation and principles of symmetry to combine the necessary conditions of good reliability with successful prediction (validity) and hopefully better theory (Wittmann 1982, 1988, Wittmann & Schmidt 1983). In my own work I was very much influenced by the London School (Spearman, Cattell, Eysenck) and ideas, concepts of Campbell/ Fiske and Brunswik. Let me first demonstrate some results in terms of the evaluation standards of Fig. 1 using my own conceptual framework.

2. Predicting behavior from Eysenck's E/N-model of personality

Eysenck's personality model with its factors extroversion and neuroticism is one of the most prominent and highly valued trait models in personality research (f.e. Amelang & Bartussek 1985). The Freiburg Personality Inventory (FPI) (Fahrenberg and Selg 1970) measures Eysenck's factors E and N. Fig. 2 shows the reanalysis of a time series study done by Fahrenberg, Myrtek, Kulick and Frommelt (1977) in terms of convergent and discriminant predictions (validity coefficients).

Fig. 2 about here

The convergent validities clearly exceed the coefficients of intelligence research. In terms of Fig. 1 the result is a "big-bang" effect bolstering Eysenck's theory to an extent rarely found in psychology. Most impressive is the whole pattern of convergent and discriminant validities.

You surely will and are obliged to ask: What kind of bad tricks have been used to produce results like these? Some answers are: No capitalization on chance, no step-wise regression, no outliers. All these usual threats to validity have been controlled and cannot be taken as an explanation. To detect possible outliers, a jackknife-procedure (Tukey 1977) was used. The 20 students were divided in ten non-overlapping subgroups of N = 2. Then one subgroup was dropped and the validities recomputed with the remaining 18 subjects which lead to ten different reanalysis. The highest convergent validity for FPIN and the neuroticism criterion was .881, the lowest .743 with a mean of .811. That part of the reanalysis was done by Schweizer (1986, p. 244). The part of the reanalysis concerning FPIE and the extraversion criterion was reported in Wittmann and Schmidt (1983). Jackknifing found the lowest convergent coefficient as .62 and the highest .75. Inspection of the respective scattergrams gave no hints to outliers.

But after assuring you of not having used bad tricks, how did we do it? Closer inspection of Fig. 2 give hints that principles of aggregation have been used, especially multiple-act-criteria in the sense of Fishbein and Ajzen (1974) and surely the phrase "Brunswik-symmetry" in the running head of my paper must have something to do with the success. Let us first have a look at a taxonomy of different kinds of criteria (see Fig. 3)

Fig. 3 about here

Fig. 3 denotes four kinds of criteria, the single-act (SAC), the repeated act (RAC), the single multiple act (SMAC) and the repeated multiple-act-criterion (RMAC). Described more deeply elsewhere (Wittmann 1982) I am convinced that broad dispositional construct like extroversion are best conceptualized as polythetic classes (Sneath and Sokal 1973). These polythetic constructs have to be validated against repeated multiple-act criteria.

RMAC combine two virtues of aggregation, first over timepoints/situations and second over different behavioral acts. The Spearman-Brown prophecy formula requisits only parallel measures. Whether these stem from measures repeated over different timepoints or different situations or different but functionally equivalent behavioral acts does not matter, reliability will improve. But according to functionally equivalent behavioral act we need theory or at least a conceptual framework what kind of different acts belong to a polythetic class.

Fortunately Eysenck gives a lot of examplesin his literature what kind of different behavioral act are tied to extroversion. The only job is to thoroughly read it through. (among many others f.e. Eysenck, 1967).
The Fahrenberg et al.(1977) study was planned to seek change-sensitive instruments. 20 students participated in a time series study over 60 days. Two times a week they appeared in the psychophysiological laboratory to take laboratory data assessment, behavior ratings and they brought with them a specially developed diary assessing moods, activation, complaints and ratings of behavioral acts. We compared this list of variables with a list of Eysenck's extroversion indicators and found 11 items to be compatible with Eysenck's. All

11 items were first z-score standardized and then summed up to a total score. The correlation of this 11 item total score with FPIE was found to be .58! The 11 item composite was only derived from theory. In psychometric theory scales are developed through item analysis, especially through principles of internal consistency. After such an analysis 3 items (number of social contacts, beer and cigarette consumption) had part-whole corrected item-total correlations lower than .20. These were excluded and the 8-variable aggregate was used as RMAC in Fig. 2. The RMAC for neuroticism was derived in a completely different way. Eysenck is not so explicit as regards behavioral acts tied to neuroticism. But from the connotation of emotional lability/stability it seems clear not to use aggregates of item means but aggregates of difference scores over a time serie to map variability, lability or changes in behavior. How can we detect items which map systematic variability? According to a multivariate reliability theory developed by the author (Wittmann, 1988) those items which had two markedly different plateaus in their respective odd-even and split-half reliability functions are top candidates. Fig. 4 demonstrates this with various mood scales.

Fig. 4 about here

In Fig. 4 the solid line reliability functions are computed the same way as Epstein (1979) did. The 60 days time series was broken down in odd and even day parts and then continuously aggregated and the two parts correlated till the maximum of two 30-day aggregates. The broken lines (split-half reliabilities) functions were computed as follows: The 60 days were broken in a set of the first 30 days and a set of the second 30 days. The days in these sets were again continuously aggregated and the two respective parts correlated up to the maximum of days 1-30 and days 31-60. My multivariate reliability theory,which is closely related to a partitioning of Cattell's data box,postulates that the odd-even functions in Fig. 4 map both systematic trait and state variance, whereas the split-half functions of Fig. 4 map less systematic state or process variance but more habitual (trait) variance.For items which demonstrated such differences in their odd-even and split-half plateaus, absolute difference scores were computed and aggregated over the whole time serie. For inclusion into the RMAC-neuroticism criterion-keying was used, i.e. only those items which had part-whole correlation with neuroticism (FPIN) greater /.20/ (see Fig.4) were included. This lead to the 7-item RMAC.

Aggregation intending to operationalize polythetic constructs is always in the danger of adding apples and oranges. Accusing someone of doing this is often used as describing violations of axioms in measurement theory or reminding one of mindless empiricism. But those critics should also pause and ask what is the question? If you intend to measure fruit as a polythetic class, there is no alternative to this kind of aggregation (Wittmann, 1982). If you want to validate such polythetic constructs you should additionally show what fruit is not, f.e. vegetables. Be sure to aggregate not only cabbage but cabbage, spinach, beans etc. to adequately map that construct. Eysenck's personality theory explicitly tells us what extroversion is not, namely neuroticism. Therefore we can use Campbell and Fiske's (1959) principles of convergent and discriminant validation to test Eysenck's theory via repeated multiple act-criteria. Fiske's (1973) question "Can a personality construct validated empirically" must be answered with a decisive yes with respect to Eysenck's theory.

3. The importance of Brunswik-symmetry

I did not tumble by trial and error over the results in Fig. 2. The basic principles and the conceptual framework is derived from Brunswik's psychology and especially from his lens model. The lens model captures the most important principle in successful science, namely symmetry. Problems and dead-ends are often overcome in natural science by principles of symmetry (Genz 1987, Zee 1986). Why should that not be true for psychology? To apply symmetry to psychology's prediction problem we must ask how our prediction models and our criteria look like with respect to symmetry. Asking questions like these it is immediately evident that our mostly hierarchical personality models need hierarchical criterion models with corresponding (symmetrical) levels of generality. Broad general second order factors like extroversion and neuroticism need broad general second order factor criteria!

Fig. 5 shows the generalized hierarchical lens model.

Fig. 5 about here

We see that only symmetrical tests at the corresponding level of generality are the fair tests of the predictive validity of a construct. Basically there is nothing wrong with asymmetry, but we must be very careful not to draw erroneous conclusions where it appears and not to expect too high correlation coefficients. Testing asymmetric relationships you should especially take care to have enough statistical power (Cohen, 1977) in your design. From a twenty year retrospect Mischel's comprehensive literature search of personality and assessment found a lot of empirical studies using asymmetric criteria, i.e. single act and narrow short-term criteria. Time series studies for the prediction problem from personality have been rarely done. No wonder that he came up with a mean of .30 which denotes nothing else than the mean of asymmetric relationships.

Brunswik-psychology is very important and fruitful for psychological research and theory testing. In my brand of it, it yields:

(a) A conceptualization via visualization (see Fig. 5).

I have extended that conceptualization not only to the classical psychometric prediction problem but also to experimental and quasiexperimental research (Wittmann, 1986, 1988) in a framework called the four-box-conceptualization. Low treatment effects may also be a function of asymmetry, i.e. a treatment too weak and too narrow for the construct it intends to change. But there is no place here for a deeper explanation.

(b) A mapping of the conceptualization into mathematics.

An elegant mathematical formalization for the lens model was given by Tucker (1964). Applied to Fig. 5 Tucker's lens model equation can be written as follows:

$$(1)\quad r_{P_S C_S} = G_{P_S C_C} R_{P_S} R_{C_S} + C_{P_S C_S} (1-R_{P_S})^{1/2} (1-R_{C_S})^{1/2}$$

$G_{P_S C_S}$ is the correlation between a linear model of the predictor and a linear model of the criterion. R_{P_S} and R_{C_S} are multiple correlation coefficients mapping the linear model of predictor and criterion respectively. The rightmost product term denotes nonlinearity and error which will be for matters of convenience abbreviated in the following as e (error). In psychometric theory we know that every true correlation is attenuated by unreliability, therefore:

$$(2)\quad G_{P_S C_S} = (r_{tt}^{P_S} \cdot r_{tt}^{C_S})^{1/2} \; G_{P_S C_S}^{true}$$

Substituting (2) in (1) using e for nonlinearity / error and appending the correction factor S for restriction of range gives:

$$(3)\ r_{P_S C_S} = S\,(r_{tt}^{P_S}\, r_{tt}^{C_S})^{1/2}\, G^{true}_{P_S C_S}\; R_{P_S}\, R_{C_S} + e$$

Factor S is equal one only when the sample variance equals the population variance. S is lower than one if the sample variance is less than population variance and S is greater than one if the sample variance is greater than the population variance. The latter is true when cases from both extremes of a distribution are overrepresented. I prefer to label R_{P_S} or R_{C_S} as coefficient of construct reliability (see Fig. 2 and Wittmann, 1988), meaning the amount of overlap (adequacy or precision) of our indicators with intended or wanted constructs.

(c) Amenability to empirical tests.

Eq.(3) allows the test of the empirical results of a series of validation studies for a construct. It is interesting to note that Eq.(3) derived from Tucker's lens model equation closely resembles the validity generalization model of Hunter & Schmidt (Hunter, Schmidt and Jackson,1982). The parameters S,$r_{tt}^{P_S}$, $r_{tt}^{C_S}$ and e are also used in their model. The extended lens model equation additionally incorporates the principles of symmetry in the parameters R_{P_S} and R_{C_S}. Two correlations may be numerically identical but refer to different levels of generality. It would be false to average them in a meta-analysis. Only theory and the respective taxonomies of generality and specifity help to avoid this trap.

4. Discussions and conclusions

The inequality $R_{P_S} > R_{C_S}$ is probably true for the majority of studies trying to predict behavior from broad dispositional variables. According to my theory this is the best explanation for Mischel's personality coefficients 20 years ago. But what about cross-situational consistency? We have aggregated over 60 days. Do we have aggregated only over time or also situations? Do we have cross-temporal stability or cross-situational consistency or do we have both? Basically time and situations are always confounded,there is no orthogonality of time and situations. There was no possibility to classify ex-post-facto the different 60 days as incorporating different situations. But few people will doubt that these 60 days did not consist of different situations. Fig. 2 tells us that consistency over time and situations is validated convergent and discriminantly for two different dispositional constructs. As regards neuroticism it must be stressed that it was not rank order consistency of arithmetic mean behavior but rank order cc sistency of variability averaged over time periods. What was consistent was the variability or inconsistency over time. Higher scores on the neuroticism trait-scale predict higher means of absolute difference scores. Whether this variability stems from situational cues or inner psychodynamic factors cannot be answered. But we see that the question for cross-situational consistent or inconsistent people has an old answer, namely those more neurotic are more variable (inconsistent) over tim and situations. The two different kinds of aggregation,i.e. z-score aggregate and difference-score aggregate remind us of an old statement by Gordon W.Allport:"Do not forget, what you decided to neglect." The kind of aggregation is very essential and should preferably be dictated by theory if you have one. If you have not, do no hesitate to use exploratory aggregation to enhance reliability, but make sure that you are doing the same kind of aggregation to your criterion. Intelligence research was successful in predicting school achievement because school achievement is also aggregated over many grades stemming from many situations, oral or written examinations,aggregated over time (i.e. one school-year) and several teachers. Intelligence research, psychology's most highly evaluated and successful contribution to science has paid attention to the parameters of eq.(3), reliability and symmetry. Remember that intelligence research declined and failed trying to predict specific criteria. Variability in all parameters of eq.(3) can be tested to explain the often disturbing variability of empirical correlation coefficients. Such analysis can clear up the picture, avoid side-tracks and dead-ends and pave the way to substantive and robust theory. Yet one of the most gratifying aspects is the evolving of continuity in psychological research. Contributions of different generations of psychologists like Spearman, Allport, Brunswik, Cattell, Campbell, Fiske, Eysenck among many others have not been futile. All have woven a common thread which will be an integral part of all future psychological tissue.

REFERENCES

Amelang,M. & Bartussek, D. (1985). Differentielle Psychologie und Persönlichkeitsforschung. 2nd.ed. Stuttgart: Kohlhammer.

Bandura,A. (1977). Social learning theory. Englewood Cliffs: Prentice Hall.

Bandura,A. (1986). Social foundations of thought and action. A social cognitive theory. Englewood Cliffs: Prentice Hall.

Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin,56,81-105.

Cohen, J. (1968). Multiple regression as a general data-analytic system. Psychological Bulletin,70,426-443.

Cohen,J. (1977). Statistical power analysis for the behavioral sciences (rev.ed.). New York: Academic Press.

Darley,J.M. & Batson,C.D.(1973). "From Jerusalem to Jericho": A study of situationa and dispositional variables in helping behaviors. Journal of Personality and Social Psychology,27,100-108.

Epstein,S. (1977). Traits are alive and well. In D.Magnusson and N.S.Endler (Eds.). Personality at the crossroads: Current issues in international psychology. Hillsdale: Erlbaum.

Epstein,S.(1979). The stability of behavior. I. On predicting most of the people much of the time. Journal of Personality and Social Psychology,37,1097-1126.

Epstein,S.(1980). The stability of behavior.II. Implications for psychological research. American Psychologist,35,790-806.

Epstein,S. (1983a). Aggregation and beyond: Some basic issues on the prediction of behavior. Journal of Personality,51,360-392.

Epstein,S.(1983b). The stability of confusion: A reply to Mischel and Peake. Psychological Review,90,179-184.

Epstein,S.(1984). The stability of behavior across time and situations.In R.A.Zucker J.Aronoff and A.I.Rabin(Eds.).,Personality and the prediction of behavior. Orlando: Academic Press.

Epstein,S.(1986). Does aggregation produce spuriously high estimates of behavior stability. Journal of Personality and Social Psychology,50,1199-1210.

Epstein,S. and O'Brien,E.J.(1985). The person-situation debate in historical and current perspective. Psychological Bulletin,98,513-537.

Eysenck,J. (1967). The biological basis of personality.Springfield:Thomas.

Fahrenberg,J. & Selg,H.(1970). Das Freiburger Persönlichkeitsinventar FPI. Göttingen:Hogrefe.

Fahrenberg,J.,Myrtek,M.,Kulick,B.& Frommelt,P.(1977). Eine psychologische Zeitreihenstudie an 20 Studenten über 8 Wochen. Archiv für Psychologie,129,242-264.

Feynman,R.(1985). Surely you're joking, Mr. Feynman! Adventures of a curious character. New York: Norton.

Fishbein,M. & Ajzen,I.(1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. Psychological Review,81,59-74.

Fiske,D.W.(1973). Can a personality construct be validated empirically.Psychological Bulletin,80,89-92.

Genz,H.(1987). Symmetrie-Bauplan der Natur. München:Piper.

Hunter,J.E., Schmidt,F.L.& Jackson,G.B.(1982). Meta-Analysis. Cumulating research findings across studies. Studying organization:Innovations in methodology (No.4). Beverly Hills:Sage.

Meehl,P.E.(1978). Theoretical risks and tabular asterisks:Sir Karl,Sir Ronald, and the slow progress of soft psychology.Journal of Consulting and Clinical Psychology,46,806-834.

Mischel,W.(1968). Personality and assessment. New York:Wiley.

Mischel,W.(1977). The interaction of person and situation. In D.Magnusson and N.S.Endler (Eds.), Personality at the crossroads. Current issues in international psychology. Hillsdale:Erlbaum.

Mischel,W. and Peake,P.K.(1982). Beyond déjà vu in the search for crossituational consistency. Psychological Review,89,730-755.

Schweizer,K.(1986). Aggregation und Vorhersagbarkeit. Vom Nutzen des Aggregationsansatzes für die Persönlichkeitsforschung. Phil.Diss.University of Freiburg. Rheinfelden:Schäuble.

Sneath,P.H. & Sokal,R.R.(1973). Numerical taxonomy. The principles and practice of numerical classification. San Francisco:Freeman.

Tucker,L.R.(1964). A suggested alternative formulation in the developments by Hursch,Hammond & Hursch; and by Hammond, Hursch & Todd. Psychological Review, 71,528-530.

Tukey,J.W.(1977). Exploratory data analysis. Reading: Addison-Wesley.

Wittmann,W.W.(1982). Der Mangel an multivariaten Betrachtungsweisen in der psychologischen Forschungspraxis mit besonderer Berücksichtigung der Interaktionismus-Debatte. In E.D. Lantermann (Ed.), Wechselwirkungen. Psychologische Analysen der Mensch-Umwelt-Beziehung. Göttingen:Hogrefe.

Wittmann,W.W.(1986). The synthesis of Cattell's BDRM,Cronbach et al. Generalizability theory and Brunswik's lens model. A framework for improving construct and predictive validity. Paper delivered San Francisco 1986: AERA Annual Meeting as part of an invited symposium:New developments in generalizability theory. International perspectives (CHair:Noreen Webb).Also available as ERIC microfiche document.

Wittmann,W.W.(1988). Multivariate reliability theory. Principles of symmetry and successful validation strategies. In J.R.Nesselroade and R.B.Cattell(Eds.), Handbook of multivariate experimental psychology 2nd ed. New York:Plenum.

Wittmann,W.W. & Schmidt,J.(1983). Die Vorhersagbarkeit des Verhaltens aus Trait-Inventaren. Theoretische Grundlagen und empirische Ergebnisse mit dem Freiburger Persönlichkeitsinventar (FPI). Research Reports No.10, Psychological Institute, University of Freiburg i.Br., West Germany.

Zee,A.(1986). Fearful symmetry. The search for beauty in modern physics. New York: Macmillan.

Fig. 1: **Evaluation standards for the empirical success/failure in psychology (correlation coefficients)**

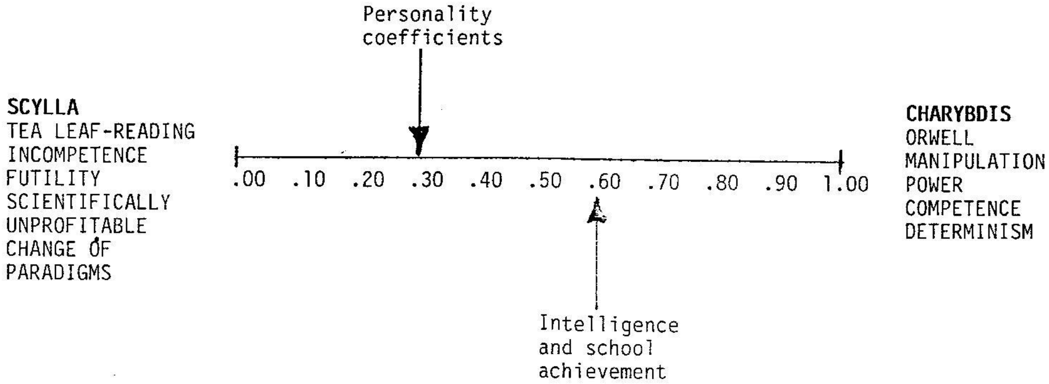
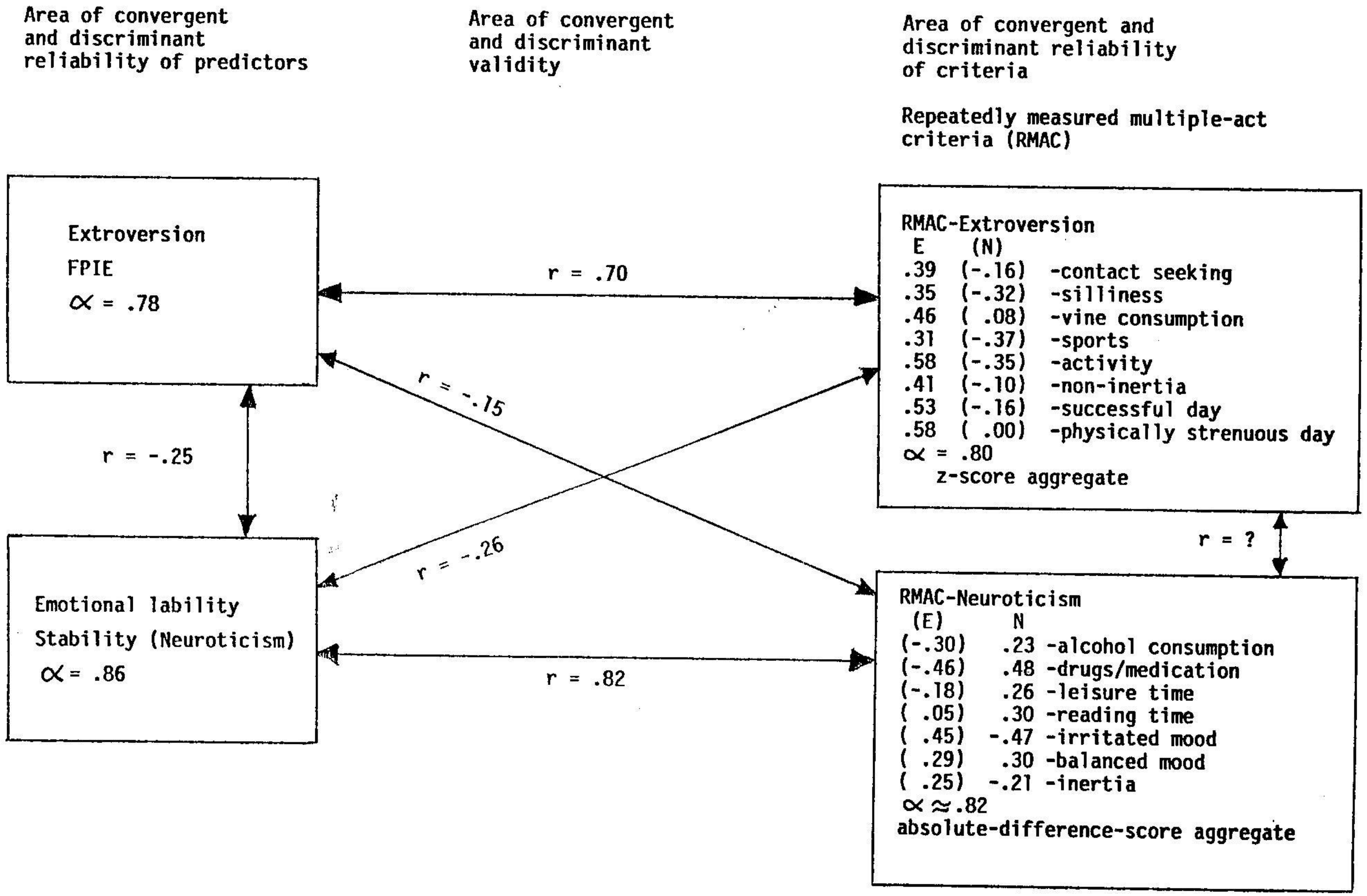


Fig. 2: TESTING EYSENCK's E/N-THEORY IN THE BRUNSWIK-SYMMETRY FRAMEWORK[a)]



a) Time series data of 20 students assessed over 8 weeks from Fahrenberg et al. 1977

Fig. 3: A TAXONOMY OF BEHAVIORAL CRITERIA

Theory directed indices ?

1 SAC

Single-act criterion
= one specific act in one specific situation
($a_{m,n}$)

2 RAC

Repeated act criterion
= repeated observation of one specific act over time and/or situations
(R_m)

3 SMAC

Single multiple act criterion
= one observation of multiple acts, belonging to one behavioral class at one time point or situation
(M_n)

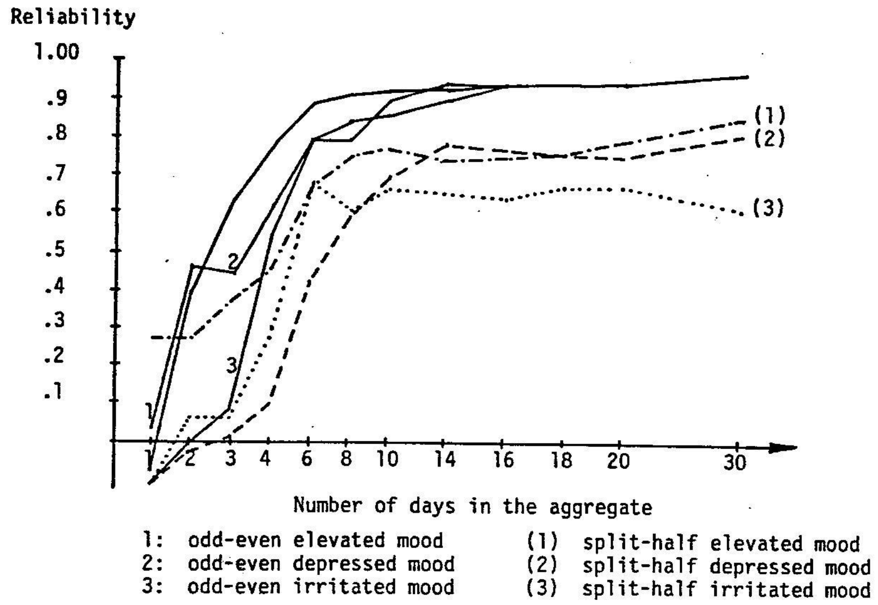
4 RMAC

Repeated multiple act criterion
= repeated observation of multiple acts belonging to one behavioral class over time and/or situations
(MR)

O b s e r v a t i o n s (Columns) → different timepoints; different situations

Behavioral acts (rows)

	1	2	3	.....	j	.....	n	
1	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$	.....	$a_{1,j}$	.....	$a_{1,n}$	$R_1 = f(a_{1.})$
2	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$	.....	$a_{2,j}$	.....	$a_{2,n}$	$R_2 = f(a_{2.})$
3	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$	.....	$a_{3,j}$	.....	$a_{3,n}$	$R_3 = f(a_{3.})$
.	.	.	.	.....	.	.....	.	.
.	.	.	.	.....	.	.....	.	.
.	.	.	.	.....	.	.....	.	.
i	$a_{i,1}$	$a_{i,2}$	$a_{i,3}$	.....	$a_{i,j}$	.....	$a_{i,n}$	$R_i = f(a_{i.})$
.	.	.	.	.....	.	.....	.	.
.	.	.	.	.....	.	.....	.	.
m	$a_{m,1}$	$a_{m,2}$	$a_{m,3}$	.....	$a_{m,j}$	.....	$a_{m,n}$	$R_m = f(a_{m.})$
	$M_1 = f(a_{.1})$	$M_2 = f(a_{.2})$	$M_3 = f(a_{.3})$		$M_j = f(a_{.j})$		$M_n = f(a_{.n})$	$MR_a = f(R_m)$
								$MR_b = f(M_n)$
								$MR_c = f(a_{m,n})$



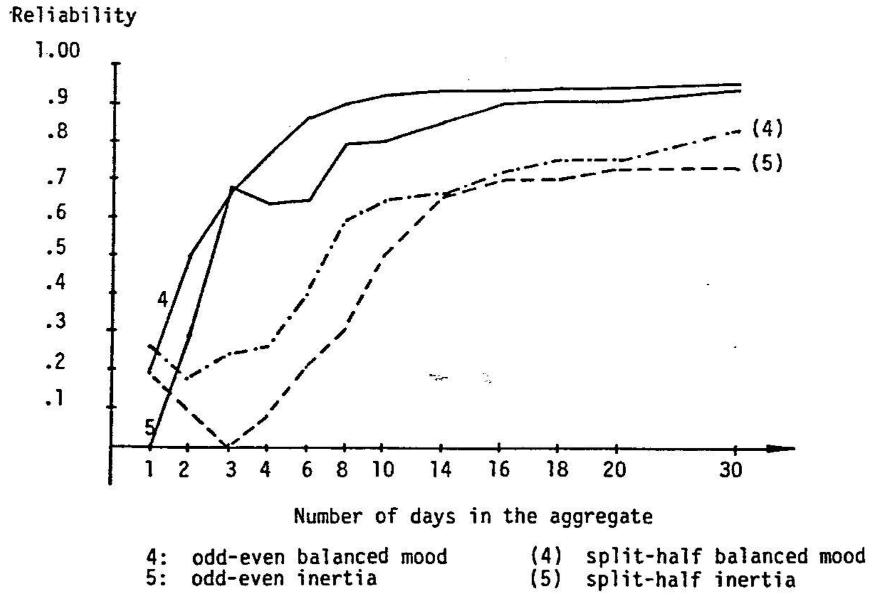


Fig. 4: Odd–even and split-half reliabilities of diverse mood scales as a function of aggregation.

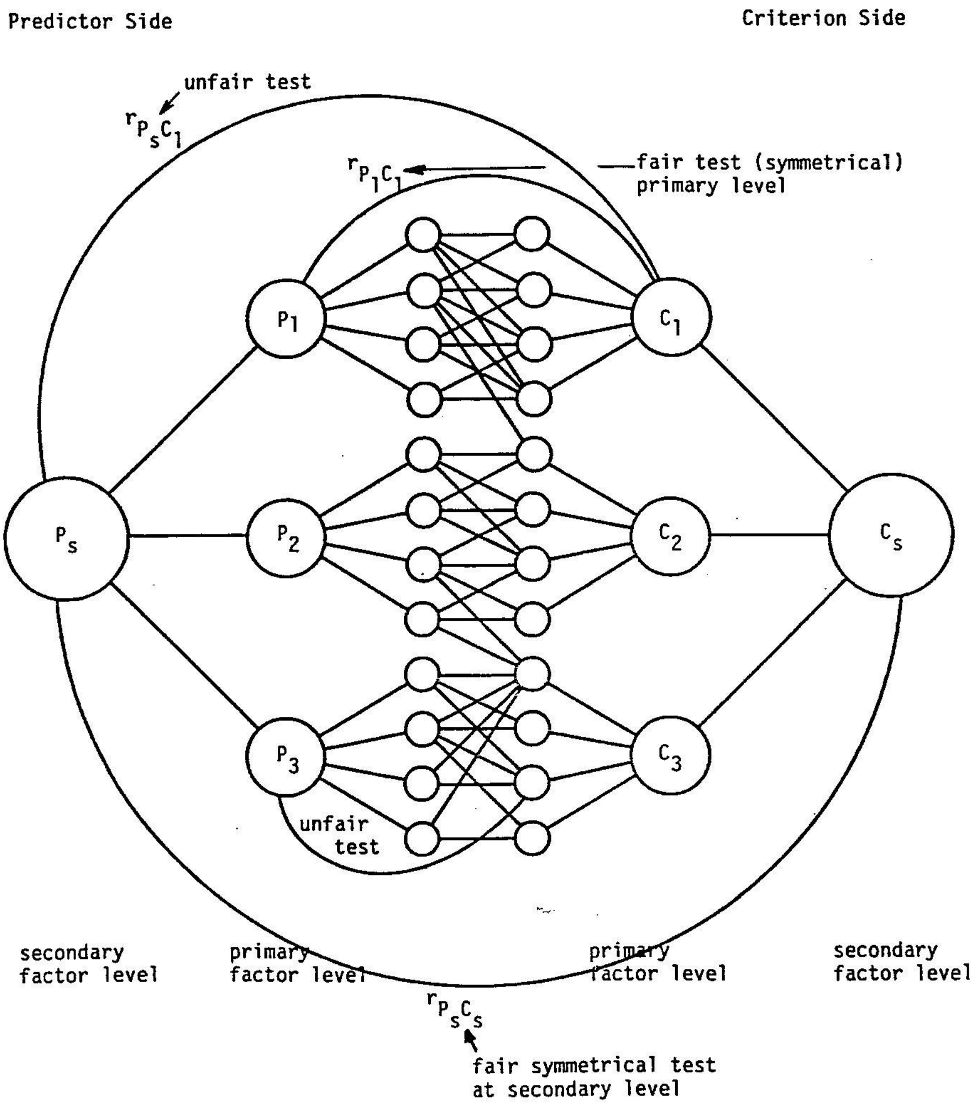


Fig. 5 : The generalized hierarchical lens model for denoting symmetry between predictors and criteria.