Effect of Dissimulation Motivation and Anxiety on Response Pattern Appropriateness Measures — Source link ⧉

Menucha Birenbaum
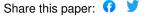
Institutions: Tel Aviv University

Related papers:

- Detecting inappropriate test scores with optimal and practical appropriateness indices

- Appropriateness measurement with polychotomous item response models and standardized indices

- Measuring the Appropriateness of Multiple-Choice Test Scores

- Caution indices based on item response theory

- The many null distributions of person fit indices

Effect of Dissimulation Motivation and Anxiety on Response Pattern Appropriateness Measures

Menucha Birenbaum
Tel-Aviv University

This study examined the effect of anxiety and dissimulation motivation of job applicants on their performance on an ability test. Two aspects of performance were considered: the total score and the appropriateness score. Four IRT-based appropriateness indices for detecting aberrant response patterns were employed in this study. The results indicate a negative effect of dissimulation motivation on the performance of low anxiety scorers, with respect to both the total score and the appropriateness score, with a greater effect on the latter. This effect was evidenced by an erratic or aberrant response pattern on the ability test; that is, missing relatively easy items while answering more difficult ones correctly. The results are discussed in light of the diverse interpretations concerning the meaning of Lie scales.

An examinee's response pattern on a test can be viewed as a product of a three-way interaction between the characteristics of the individual, of the test, and of the situation in which the test is given. One personality trait whose effect on test performance has been thoroughly investigated is anxiety (e.g., Hamilton, 1985; Houston, 1982; Sarason, 1984; Sarason & Spielberger, 1982). Difficulty in concentrating is indicated as one of the diagnostic criteria for generalized anxiety disorders (cf. American Psychiatric Association, 1980, p. 233). The cognitive approach to anxiety (i.e., the informa-

tion-processing view which states that anxiety arises from a self-assessment of personal deficit in meeting situational demands) has helped to clarify the concept of anxiety and the ways in which it affects performance (Sarason, 1984). Hamilton (1985) explained that "for high anxious individuals anxiety is cognitive information with low retrieval thresholds and high priorities for directing attention to stimuli which have no specific relationship to the task" (p. 18). According to Sarason (1978), evaluating situations increase the interfering thoughts of highly anxious individuals, thus diminishing attention to the task and interfering with its efficient execution.

Special attention is given in the psychological and educational literature to test anxiety (e.g., Geen, 1976, 1980; Sarason, 1975, 1980; Spielberger, Anton, & Bedell, 1976; Spielberger, Gonzales, Taylor, Algaze, & Anton, 1978; Tryon, 1980; Wine, 1971, 1980). Yet most studies regarding the effect of anxiety on test performance focus primarily on an examinee's total test score. Only a few investigators have studied the effect of anxiety on the "quality" or the appropriateness of an examinee's response pattern on a test. Schmitt and Crocker (1984), who examined this aspect, reported significant correlations between test anxiety and several person fit measures. On the basis of findings from several studies concerning the effect of anxiety on cognitive task performance, Eysenck (1979, 1984) concluded that anxiety has a greater effect

on processing efficiency than on performance effectiveness.

The effect of the situation in which the test is given has also been widely discussed in the psychological literature. The tendency of examinees to "fake good" when examined under selection settings has long been acknowledged. Different methods to cope with such distortions have been suggested. These include special scales for measuring dissimulation motivation ("Lie" scales), as well as techniques for reducing the distortion by causing examinees to believe that their honesty can be checked.

Eysenck and Eysenck (1976) summarized available interpretations for high Lie scores. The most common is the desire to deliberately impress or deceive the assessor. Another is the tendency for examinees to respond to an idealized version of his or her behavior, rather than to the real self. Yet another interpretation posits a mistaken but genuine belief on the part of the examinee that he or she is indeed being truthful. The last interpretation, which is the most controversial, considers a high score on a Lie scale to be the response given by a "genuine" conforming person who sincerely accepts social regulations and norms. Supporters of this interpretation stress its consistency with the way all other personality scales are interpreted (Allsopp & Feldman, 1976; Branthwaite, Trueman, & Berrisford, 1981; Emler, 1984; O'Hagan & Edmunds, 1982).

Eysenck and Eysenck (1975) stated that "the Lie scale may measure different constructs depending on the subject's motivation to dissimulate. While it does measure dissimulation under dissimulation conditions, in conditions of low dissimulation the score may be used to measure the personality function underlying the scale" (p. 15). Michaelis and Eysenck (1971) claimed that a negative correlation ($> -.5$) between Neuroticism and Lie scores provides a fairly adequate index of the existence of dissimulation motivating conditions. The susceptibility of anxiety scales to dissimulation has been noted by others as well (e.g., Cattell, Eber & Tatsuoka, 1970; McCrae & Costa, 1983).

If the first and most common explanation of high Lie scores holds and Eysenck's distinction between dissimulation motivating conditions and conditions of non-motivation is correct, then it can be hypothesized that in a selection situation, the performance of low anxiety scorers on an ability test will be affected by their Lie scores. Stated more specifically, a sample of examinees with low anxiety scores but with high Lie scores will exhibit (as will high anxiety scorers) a less appropriate response pattern on an ability test than low-anxiety examinees who scored low on a Lie scale. The present study investigates this hypothesis.

Several indices for detecting aberrant or inappropriate response patterns have been developed in recent years (cf. Donlon & Fischer, 1968; Levine & Rubin, 1979; Sato, 1975; Tatsuoka & Linn, 1983; Tatsuoka & Tatsuoka, 1982, 1983; van der Flier, 1977, 1982; Wright & Stone, 1979). These indices can be classified into two groups. One group consists of indices based on actual observed response patterns, such as Sato's caution index, van der Flier's U index, Donlon and Fischer's personal biserial, and Tatsuoka and Tatsuoka's norm conformity index (NCI). (For a summary of these indices see Cliff, 1982, and Harnisch & Linn, 1981). The other group consists of indices based on item response theory (IRT) stochastic models (Lord & Novick, 1968). Indices classified into this group are the fit statistics developed by Wright and his associates, the appropriateness indices developed by Levine and his associates (Levine and Rubin, 1979), and the group of extended caution indices (ECI) developed by Tatsuoka and her associates. (For a summary of these indices see Harnisch & Tatsuoka, 1983).

Method

Examinees

The sample consisted of 724 male Israeli applicants for various driving jobs, who were administered the MMPI and the 28-item Scale B as part of a selection program. The mean age of the examinees was 28.11 years, with a standard deviation of 7.42 years. The mean education of the sample was 11.2 years of schooling, with a standard deviation of 3.8 years.

Instruments

The Lie (L) scale of the MMPI (Hathaway & McKinley, 1967) was used as a measure of dissimulation motivation. The Lie scale includes 15 items that were selected on a rational basis to identify persons who are deliberately trying to "fake good" on the questionnaire. The items refer to minor social faults that most people are generally willing to admit. The majority of the normative group answered 12 of the items in the nondeviant direction. In Israeli normative samples, the modal response on the L scale ranged from 4 to 6 (Montag, 1978).

The test-retest reliability coefficients for intervals up to 1 week on the L scale in American samples ranged from .70 to .85 (Greene, 1980). Scores on the Hebrew version of the Lie scale were found to correlate .54 with scale R (Response Bias) of the Comrey Personality Scales (Montag & Comrey, 1982), and .47 with the Lie scale of the Eysenck Personality Questionnaire.

The Psychasthenia (Pt) scale of the MMPI (Hathaway & McKinley, 1967) was used as a measure of anxiety. Pt is a 48-item scale designed to assess a neurotic syndrome. According to Greene (1980), "The anxiety assessed by this scale is of a long term nature or trait anxiety although the scale is somewhat responsive to situational stress as well" (p. 99). Comrey (1958) identified seven principal factors underlying responses to this scale: neuroticism, anxiety, withdrawal, poor concentration, agitation, psychotic tendencies, and poor physical health. Scores on the Hebrew version of the Pt scale were found to correlate .48 with scale N (Neuroticism) of the Eysenck Personality Questionnaire and $-.53$ with scale S (Stability) of the Comrey Personality Scales (Montag & Comrey, 1982).

According to Greene (1980), test-retest reliability coefficients for the Pt scale range from .74 to .93 in intervals up to 2 weeks. These coefficients are higher than those of the other clinical scales of the MMPI.

Scale B (Intelligence) of the 16PF questionnaire (Cattell et al., 1970), a multiple-choice test, was used as a measure of general ability. According to Cattell et al. (1970), "the aim in constructing the B measure has been to keep a balance between emphasis on the fluid and crystalized general ability factors" (p. 82). (For empirical findings concerning the validity of the scale see also Fleishman & Fine, 1971.) In the Hebrew version of the 16PF questionnaire, Scale B is administered as a separate test. The version of Scale B used in the present study included 28 items from forms A and B of the 16PF and the Clinical Analysis Questionnaire. The Cronbach alpha (α) coefficients of internal consistency for the Hebrew version of the scale range from .76 to .85 in various samples (cf. Zak, 1976).

Appropriateness Indices

Four appropriateness indices based on IRT models were used (Tatsuoka's standardized extended caution indices and Levine's standardized appropriateness index). The standardized indices were preferred over the unstandardized ones because the latter were found to provide inflated values at both extremely high and low total scores (cf. Drasgow, 1982; Harnisch & Tatsuoka, 1983).

The group of ECIs developed by Tatsuoka (1982, 1984; Tatsuoka & Linn, 1983) are defined as the complement of the ratio of two covariances. These are extensions of Sato's (1975) caution index based on sample statistics, with a probability matrix substituting for the observed binary matrix. Sato's data matrix is a students $\times$ items (row $\times$ column) matrix in which the students have been arranged in descending order of their total scores and the items in ascending order of difficulty. The probability matrix (P_{ij}) used by Tatsuoka and Linn replaces the matrix of binary scores of n items for N examinees (Y_{ij}) by the probability of examinee i answering item j correctly. These probabilities may be based on the one-, two-, or three-parameter logistic model. Tatsuoka and Linn showed analogous relationships between student (S) curves and problem (P) curves used by Sato, and test response curves (TRC) and group response curves (GRC) respectively, in logistic models.

Sato's caution index is defined as the complement of the ratio of two covariances. The numerator is the covariance of an observed response vector and the column-sum vector of the data matrix

(Y_{ij}). The denominator is the covariance of a Gutt-man vector with the same total score as the observed vector and the column-sum vector. The first standardized extended caution index ($ECI1_z$) replaces the Guttman vector with the ith row vector of the probability matrix. $ECI2_z$ replaces the column-sum vector in the observed binary matrix with the GRC vector which is the average of the column-sum vector of the probability matrix. $ECI4_z$ replaces the GRC vector in the numerator of $ECI2_z$ with the probability vector from the ith row in the probability matrix. The lower the score on these indices, the higher the appropriateness of the response pattern.

The appropriateness measures suggested by Levine and his associates (Drasgow, 1982; Levine & Rubin, 1979) are based on the notion that an appropriate response pattern is one which is representative of the group whose abilities are measured by the test. Since maximum-likelihood procedures are used for estimating the item parameters and the examinees' abilities, an inappropriate response pattern can be defined as one that does not contribute much to maximizing the likelihood function. The appropriateness index, L_z, is the standardized measure of the examinee's contribution to the likelihood function. The higher the score on this index, the more appropriate the response pattern.

Analysis

Item parameters for Scale B based on the two-parameter logistic model were computed using the computer program LOGIST (Wingersky, Barton, & Lord, 1982) for a sample of 1,237 Israeli males. A FORTRAN TCC3 program (Baillie & Tatsuoka, 1982) was employed for computing the four appropriateness indices for the subgroup of 724 examinees who completed all three questionnaires.

Examinees in the lower and higher thirds of the Pt distribution (scores ≤ 7 or ≥ 14) were considered low and high anxiety scorers respectively. Of these groups, those who scored among the highest third on the distribution of the L scale (scores ≥ 6) were classified into the low anxiety-high Lie (LAHL) and high anxiety-high Lie (HAHL) groups respectively. Those who scored in the lower third of the L distribution (scores ≤ 3) were classified into the low anxiety-low Lie (LALL) and high anxiety-low Lie (HALL) groups respectively. The LALL, LAHL, HALL and HAHL groups consisted of 39, 173, 112, and 32 examinees respectively.

Results

The correlations among the variables for the entire sample are presented in Table 1. As can be seen, all the correlation coefficients are significant at the .01 level, but they vary considerably with respect to their magnitude. The four appropriateness indices are highly correlated among themselves. The correlation between the Lie and Pt scales is negative and relatively moderate; so are the correlations between the score on Scale B and the appropriateness measures. Both the Lie and the Pt scales yielded relatively low negative correlations

Table 1
The Correlation Matrix Between the Four
Appropriateness Indices, Scores on Scales
L, Pt, B and on $\hat{\theta}$(B) (N=724)

	Pt	B	$\hat{\theta}$(B)	$ECI1_z$	$ECI2_z$	$ECI4_z$	L_z
L	-.49	-.14	-.14	.11	.11	.11	-.11
Pt		-.15	-.12	.13	.14	.14	-.14
B			.93	-.55	-.55	-.49	.57
$\hat{\theta}$(B)				-.38	-.40	-.33	.41
$ECI1_z$					.98	.97	-.99
$ECI2_z$						.98	-.99
$ECI4_z$							-.97

Note. All correlations are significant at .01 level.

with the appropriateness measures and with the total score on Scale B. (Note that L_z is oriented opposite to the ECIs.)

The means, standard deviations and F ratios of the one-way ANOVA for the four groups on the performance scores for Scale B are presented in Table 2. As can be seen, the LALL group scored significantly higher on Scale B and exhibited more appropriate response patterns than did the other three groups, which did not differ significantly from each other, as indicated by Tukey's multiple-range test. In the case of the ability estimate $\hat{\theta}(B)$ and the total test score, there is a second significant difference between the HAHL group, which scored the lowest, and the other three groups.

In order to illustrate the meaning of the appropriateness scores, the test items were ordered according to their difficulty level (proportion correct) in the entire sample and in each of the two groups of main interest in the present study (i.e., low anx-

iety scorers with low and high Lie scores). The proportions correct and *t*-test values are presented in Table 3. There are significant differences between the two groups on the four easiest items, whereas the differences for the six most difficult items are insignificant. Two-thirds of the other items yielded insignificant differences between the two groups.

Discussion

The negative correlation between Lie and Pt scores in the present study indicates, according to the criterion suggested by Michaelis and Eysenck (1971), the existence of dissimulation-motivating conditions under which the examinees responded to the questionnaires. Since anxiety is particularly susceptible to dissimulation, it is reasonable to anticipate that, in addition to genuine low-anxiety examinees, the group of low anxiety scorers includes

Table 2

One-Way ANOVA Results for the Differences Among the 4 Groups on the Total Score and the Appropriateness Scores on Scale B

Index and Stat.	Group Means				ANOVA		Multiple Range Test[b]
	LALL	LAHL	HALL	HAHL	F[a]	sig.	
ECI1$_z$							
Mean	−.06	1.03	1.03	1.51	7.48	**	(1),(2,3,4)
SD	1.22	1.63	1.41	1.55			
ECI2$_z$							
Mean	−.05	1.21	1.24	1.82	9.59	**	(1),(2,3,4)
SD	1.20	1.70	1.46	1.71			
ECI4$_z$							
Mean	−.12	1.14	1.21	1.66	9.83	**	(1),(2,3,4)
SD	1.21	1.65	1.39	1.67			
L_z							
Mean	.11	−1.27	−1.34	−1.92	9.42	**	(1),(4,3,2)
SD	1.25	1.88	1.67	1.82			
$\hat{\theta}(B)$							
Mean	1.00	.62	.61	.22	9.40	**	(1),(4,3,2);(4),(3,2,1)
SD	.50	.65	.61	.55			
TOT(B)							
Mean	16.79	14.24	14.09	11.47	12.53	**	(1),(4,3,2);(4),(3,2,1)
SD	3.27	3.76	3.83	2.83			

**p < .001
[a]df:3,352
[b]Significant differences for Tukey's multiple range test (cv ∝.05= 3.66).

Table 3
Item Difficulties for the Total Sample
and the LALL LAHL Groups

Item	Proportion Correct Total	LALL	LAHL	t value	df	sig.
5	.90	1.00	.90	2.05	210	*
4	.90	.97	.88	2.54	117	*
16	.88	.97	.91	1.98	103	*
1	.85	.95	.82	2.77	96	**
2	.82	.90	.81	1.31	210	NS
13	.80	.82	.82	.08	210	NS
22	.75	.85	.72	1.60	210	NS
24	.74	.82	.75	.92	210	NS
9	.74	.87	.73	1.82	210	NS
19	.68	.95	.71	4.88	120	**
20	.66	.72	.64	.90	210	NS
23	.64	.82	.63	2.29	210	*
25	.63	.85	.60	3.53	73	**
12	.61	.85	.59	3.69	73	**
26	.58	.67	.59	.89	210	NS
14	.58	.69	.60	1.06	210	NS
21	.57	.77	.61	1.91	210	NS
6	.52	.79	.54	2.99	210	**
8	.50	.59	.51	.85	210	NS
18	.39	.33	.41	-.89	210	NS
15	.37	.41	.40	.13	210	NS
3	.36	.54	.34	2.39	210	*
28	.35	.21	.31	-1.33	210	NS
27	.31	.38	.29	1.09	210	NS
11	.25	.26	.28	-.27	210	NS
10	.24	.18	.20	-.32	210	NS
7	.21	.15	.23	-.99	210	NS
17	.20	.31	.17	1.91	210	NS

*p < .05
**p < .01

This interpretation leads to the conclusion that an examinee may be able to falsify his or her responses on an anxiety test, but not the effect of anxiety on his or her cognitive functioning. This effect is expressed in an erratic or aberrant response pattern on the ability test (i.e., missing relatively easy items while answering more difficult ones correctly), as can be seen from the comparison of item difficulties in the two groups. The attentional interpretation of anxiety provided by the cognitive approach helps to clarify this effect, which is especially noticeable in evaluating situations. According to Geen (1976, 1980) and Sarason (1984), it is the worry over evaluation that leads to task-irrelevant cognitions that interfere with attention to the range of cues in the situation.

Supporters of the competing explanation concerning the meaning of Lie scales may argue that it is the compulsiveness aspect, characteristic of a high-conforming person, that is responsible for the aberrant response patterns in the low anxiety-high Lie group. According to Cattell et al. (1970), one of the main contributors to the second order factor of conformity is compulsiveness (Q3), which is described as "the extent to which the person has crystallized for himself a clear, consistent, admired pattern of socially approved behavior, to which he makes definite efforts to conform" (p. 107). Gordon (1985) proposed a theory which predicts that obsessive-compulsive disorder is characterized by "hyperattention and that under mild stress these subjects would therefore show disruption of attention, particularly in controlled information processing" (p. 105). When stress is increased this would cause, according to Gordon, "greater deterioration in performance in these subjects than in other groups, seen in reduced accuracy and slower reaction " (p. 105). The fact that the high anxiety-high Lie group scored the highest on the inappropriateness measures, and lowest on the total and ability scores, seems to support this interpretation.

Taken together, it seems that both anxiety and compulsiveness contribute to the inappropriateness of the responses. Anxiety is particularly dominant in the high anxiety-low Lie group, compulsiveness in the low anxiety-high Lie group, and both in the high anxiety-high Lie group. The fact that anxiety

some high-anxiety examinees who tried to impress or deceive their assessors by reporting a low anxiety level in order to get the jobs for which they applied. This interpretation, which follows the most common explanation of the meaning of the Lie score, seems to have gained substantial support from the results of the present study. There was a significant difference in ability performance between low and high Lie scorers in the low-anxiety group. Moreover, there was no significant difference in ability performance between the latter group and the high-anxiety groups.

and compulsiveness affect performance, due to attentional disturbances which result in missing easy items on the ability test, should be taken into consideration when assessing the ability of examinees exhibiting those personality characteristics. The implications for test practitioners are quite obvious. In assessing an examinee's ability in an evaluating situation, the appropriateness of the response pattern must be considered as well as the total test score.

The four indices used in the present study for measuring response pattern appropriateness yielded similar results. Their high intercorrelations confirm previous results (Harnisch & Tatsuoka, 1983; Rudner, 1983; Birenbaum, 1985). They can, therefore, be recommended as alternate measures of response pattern appropriateness.

References

Allsopp, J., & Feldman, M. P. (1976). Personality and antisocial behavior in school boys. *British Journal of Criminology, 16,* 337–351.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.) Washington DC: Author.

Baillie, R., & Tatsuoka, K. K. (1982). TCC3 [Computer program]. Urbana IL: University of Illinois at Urbana-Champaign, Computer-based Education Research Laboratory.

Birenbaum, M. (1985). Comparing the effectiveness of several IRT based appropriateness measures in detecting unusual response patterns. *Educational and Psychological Measurement, 45,* 523–534.

Branthwaite, A., Trueman, M., & Berrisford, T. (1981). Unreliability and marking: Further evidence and possible explanation. *Educational Review, 33,* 41–46.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the 16PF.* Champaign IL: Institute for Personality and Ability Testing.

Cliff, N. (1982). *Evaluating Guttman scales, 1941–81: Some old and new theories.* Paper presented at the American Psychological Convention, San Francisco.

Comrey, A. L. (1958). A factor analysis of items on the MMPI psychasthenia scale. *Educational and Psychological Measurement, 18,* 293–300.

Donlon, T. F., & Fischer, F. F. (1968). An index of an individual's agreement with group-determined item difficulties. *Educational and Psychological Measurement, 28,* 105–113.

Drasgow, F. (1982). Choice of test model for appropriateness measurement. *Applied Psychological Mea-*
surement, 6, 297–308.

Emler, N. (1984). Differential involvement in delinquency: Toward an interpretation in terms of reputation management. *Progress in Experimental Personality Research, 13,* 173–239.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire (junior and adult).* London: Hodder & Stoughton.

Eysenck, H. J., & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality.* London: Hodder & Stoughton.

Eysenck, M. (1979). Anxiety, learning and memory: A reconceptualization. *Journal of Research in Personality, 13,* 363–385.

Eysenck, M. W. (1984). *Trait anxiety and cognitive task performance.* Paper presented at the NATO Advanced Study Institute, Athens, Greece.

Fleishman, J. J., & Fine, B. J. (1971). Note on cognitive factors related to factor B of the 16PF test. *Psychological Reports, 29,* 1075–1077.

Geen, R. G. (1976). Test anxiety, observation and range of cue utilization. *British Journal of Social and Clinical Psychology, 15,* 253–259.

Geen, R. G. (1980). Test anxiety and cue utilization. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 43–61). Hillsdale NJ: Erlbaum.

Gordon, P. K. (1985). Allocation of attention in obsessional disorder. *British Journal of Clinical Psychology, 24,* 101–107.

Greene, R. L. (1980). *The MMPI—an interpretive manual.* New York: Grune & Stratton.

Hamilton, V. (1985). A cognitive model of anxiety; implications for theories of personality and motivation. In C. D. Spielberger, I. G. Sarason, & P. B. Defares (Eds.), *Stress and anxiety* (vol. 9, pp. 17–35). Washington DC: Hemisphere.

Harnisch, D. L., & Linn, R. L. (1981). *Identification of aberrant response patterns* (Final Report for Grant No. G–9–0003). Washington DC: National Institute of Education.

Harnisch, D. L., & Tatsuoka, K. K. (1983). A comparison of appropriateness indices based on item response theory. In R. Hambleton (Ed.), *Applications of item response theory.* Vancouver: Educational Research Institute of British Columbia.

Hathaway, S. R., & McKinley, J. C. (1967). *MMPI manual* (rev. ed.). New York: The Psychological Corporation.

Houston, K. (1982). Trait anxiety and cognitive coping behavior. In I. G. Sarason & C. D. Spielberger (Eds.), *Achievement, stress, and anxiety.* Washington DC: Hemisphere.

Levine, M. V., & Rubin, D. B. (1979). Measuring the appropriateness of multiple choice test score. *Journal of Educational Statistics, 4,* 269–290.

Lord, F. M., & Novick, M. R. (1968). *Statistical the-*

ories of mental test scores. Reading MA: Addison-Wesley.

McCrae, R. R., & Costa, P. T., Jr. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51*, 882–888.

Michaelis, W., & Eysenck, H. J. (1971). The determination of personality inventory factor patterns and intercorrelations by changes in real life motivation. *Journal of Genetic Psychology, 118*, 223–234.

Montag, I. (1978). *Adequate response set and social desirability in personality questionnaires*. Unpublished doctoral dissertation, Tel-Aviv University.

Montag, I., & Comrey, A. L. (1982). Comparison of certain MMPI, Eysenck and Comrey personality constructs. *Multivariate Behavioral Research, 17*, 93–97.

O'Hagan, F. J., & Edmunds, G. (1982). Teachers' observations on pupils' untruthfulness in relation to the "Lie" scale. *Personality and Individual Differences, 3*, 335–338.

Rudner, L. M. (1983). Individual assessment accuracy. *Journal of Educational Measurement, 20*, 207–219.

Sarason, I. G. (1975). Test anxiety and cognitive modeling. *Journal of Personality and Social Psychology, 43*, 148–153.

Sarason, I. G. (1978). The test anxiety scale: Concept and research. In C. D. Spielberger and I. G. Sarason (Eds.), *Stress and anxiety* (vol. 5, pp. 193–216). Washington DC: Hemisphere.

Sarason, I. G. (Ed.). (1980). *Test anxiety: Theory, research and applications*. Hillsdale NJ: Erlbaum.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology, 46*, 929–938.

Sarason, I. G., & Spielberger, C. D. (1982). *Achievement, stress and anxiety*. Washington DC: Hemisphere.

Sato, T. (1975). *The construction and interpretation of S-P tables*. Tokyo: Meiji Tosho. (In Japanese.)

Schmitt, A. P., & Crocker, L. (1984). *The relationship between test anxiety and person fit measures*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Spielberger, C. D., Anton, W. D., & Bedell, J. (1976). The nature and treatment of test anxiety. In M. Zuckerman & C. D. Spielberger (Eds.), *Emotions and anxiety: New concepts, methods and applications*. Hillsdale NJ: Erlbaum.

Spielberger, C. D., Gonzales, H. P., Taylor, C. J., Algaze, B., & Anton, W. D. (1978). Examination stress and test anxiety. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and anxiety* (vol. 5). New York: Hemisphere/Wiley.

Tatsuoka, K. K. (1982). *Standardization extended caution indices and comparison of their error detection rates* (Research Report No. 82-4-ONR). Urbana IL:

University of Illinois, Computer-based Education Research Laboratory.

Tatsuoka, K. K. (1984). Caution indices based on item response theory. *Psychometrika, 49*, 95–110.

Tatsuoka, K. K., & Linn, R. L. (1983). Indices for detecting unusual patterns: Links between two general approaches and potential applications. *Applied Psychological Measurement, 7*, 81–96.

Tatsuoka, K. K., & Tatsuoka, M. M. (1982). Detection of aberrant response patterns and their effect on dimensionality. *Journal of Educational Statistics, 7*, 215–231.

Tatsuoka, K. K., & Tatsuoka, M. M. (1983). Spotting erroneous rules of operation by individual consistency index. *Journal of Educational Measurement, 20*, 221–230.

Tryon, G. S. (1980). The measurement and treatment of test anxiety. *Review of Educational Research, 50*, 343–372.

van der Flier, H. (1977). Environmental factors and deviant response patterns. In Y. H. Poortinga (Ed.), *Basic problems in cross cultural psychology*. Amsterdam: Swets & Seitlinger.

van der Flier, H. (1982). Deviant response patterns and comparability of test scores. *Journal of Cross-Cultural Psychology, 13*, 267–298.

Wine, J. D. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76*, 92–104.

Wine, J. D. (1980). Cognitive-attentional theory of test anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research and applications*. Hillsdale NJ: Erlbaum.

Wingersky, M. S., Barton, M. A., & Lord, F. M. (1982). *LOGIST user's guide*. Princeton NJ: Educational Testing Service.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: Mesa Press.

Zak, I. (1976). *Psychometric properties of the 16PF Questionnaire*. Tel-Aviv: Ramot. (In Hebrew.)

Acknowledgments

Author's Address

Send requests for reprints or further information to Menucha Birenbaum, University of Illinois at Urbana-Champaign, CERL, 252 ERL, 103 S. Mathews St., Urbana IL 61801 (before September 1, 1986); Menucha Birenbaum, Unit for Evaluation Research and Measurement, School of Education, Tel-Aviv University, Ramat-Aviv 69978, Tel-Aviv, Israel (after September 1, 1986).