

**EUROPEAN
PERSPECTIVES
IN PSYCHOLOGY**

A Three Volume Work

VOLUME 1

*Theoretical Psychometrics Personality Developmental
Educational Cognitive Gerontological*

VOLUME 2

*Clinical Health Stress and Anxiety Neuropsychology
Psychophysiology*

VOLUME 3

Work and Organizational Social and Economic Cross-cultural

**EUROPEAN
PERSPECTIVES
IN PSYCHOLOGY**

VOLUME 1

*Theoretical Psychometrics Personality Developmental
Educational Cognitive Gerontological*

Edited by

**PIETER J. D. DRENTH
JOSEPH A. SERGEANT
ROELF J. TAKENS**

JOHN WILEY & SONS

Chichester · New York · Brisbane · Toronto · Singapore

Further evidence on the invalidity of self-reported consistency

Manfred Schmitt

Introduction

The general trait model of personality relies on an assumption of behavioural consistency. As dispositional constructs, traits are useful for descriptive and predictive purposes only if individual differences in behaviour remain fairly stable across different situations, modes of responding, and times of measurement. Mischel (1968) reviewed empirical research pertaining to the relative consistency of behaviour and concluded that the trait model might be appropriate for a rough description of individual differences in broad personality dimensions, but that it was not useful for the prediction of specific behaviour. During the past 20 years, Mischel's critique triggered a debate by a number of scholars about the validity and usefulness of the trait model. From the many theoretical, methodological, and empirical reactions to Mischel's review, the most promising direction appears to involve a refinement of the consistency assumption (e.g., Amelang, Kobelt, & Frasch, 1985; Emmons & Diener, 1986; Fazio & Zanna, 1981; Monson, Hesley, & Chernick, 1982). This refinement is based on the idea that relative consistency may not occur across all individuals, all situations, all times, and all behavioural modes but within restricted ranges of these dimensions or perhaps only within combinations of these restricted ranges (c.f., Asendorpf, 1988).

One of the most prominent attempts to refine the consistency assumption was made by Bem and Allen (1974). They measured individual differences in behavioural consistency using people's self-report of consistency in their behaviour. In accord with their theoretical expectations, Bem and Allen (1974) found an average correlation of .57 among various indicators of friendliness for subjects who rated themselves to be consistently friendly. By contrast, the average correlation among indicators was only .27 for subjects

who claimed that their friendliness varied across different situations. Self-reported consistency for a second trait, conscientiousness, did not show a similar moderator effect, however. Moreover, in most subsequent research, there has been either a failure to replicate Bem and Allen's finding (e.g., Amelang & Borkenau, 1981; Borkenau, 1981; Chaplin & Goldberg, 1984; Lippa & Mash, 1981; Mischel & Peake, 1982; Paunonen & Jackson, 1985; Wymer & Penner, 1985) or findings have been equivocal due to methodological problems (c.f., Paunonen & Jackson, 1985; Rushton, Jackson, & Paunonen, 1981). In two of the most recent investigations of the validity of self-reported consistency, more than 300 students and their roommates rated the subject's personality on ten (Zuckerman, Bernieri, Koestner, & Rosenthal, 1989) or eleven (Zuckerman, Koestner, DeBoy, Garcia, Maresca, & Sartoris, 1988) trait dimensions. In addition, subjects either rated or ranked these traits with respect to cross-situational consistency in behaviour. The direction of most moderator effects of consistency on the self-peer correlations was consistent with expectations. Even so, effect sizes were quite small, and only one out of 10 (Zuckerman et al. 1989) or two out of 11 (Zuckerman et al. 1988) were statistically significant.

The apparent failure of many of these studies to demonstrate the moderating effects of self-reported consistency cannot be interpreted as firm evidence for the invalidity of self-report measures of behavioural consistency. Two major problems preclude this interpretation. First, the *reliabilities* of the trait-specific consistency-ratings suggested by Bem and Allen (1974) are typically quite low, usually below .30 (Amelang et al., 1985; Borkenau, 1981; Greaner & Penner, 1982). Because of the low reliability of self-report measures of consistency, large validity coefficients cannot be expected. Second, the self-report measure of consistency developed by Bem and Allen (1974) and used by most other researchers entails subjects' evaluations of *variability* in their behaviour across different situations. However, because the situations are not specified for a subject, he or she must select which situations to use for estimating variability in his or her behaviour. Consequently, it is likely that different subjects refer to different situations in providing evaluations of how variable their behaviour is for the trait in question. Thus, individual differences in the measure devised by Bem and Allen (1974) do not necessarily reflect person by situation interactions. Conceivably, they may simply reflect individual differences in the range of difficulty of the situations that different subjects select for comparison. Because of these limitations, it would be more appropriate to have subjects either apply the concept of correlation (or effect size) when rating their consistency or to let

them judge the stability of a given (latent) trait itself across time rather than the consistency of its (manifest) indicators.

Method

Dimensions of consistency: A new self-report measure

To address these issues, we developed a measure of consistency in the domain of attitudes (Schmitt, Dalbert, & Montada, 1982). The measure is a self-report questionnaire of consistency that contains three moderately correlated scales (c.f., Schmitt, 1988)

The first scale is *Centrality of Consistency*. *Centrality* involves the tendency of a person to form his or her own attitudes, the motivation to behave in accordance with these attitudes, as well as the perception of actually doing so. Thus, subjects with high *Centrality* would tend to agree with the following items: (a) It is important for me to comply with my attitudes; (b) It is crucial for me to develop my own point of view; (c) I often stand up for my beliefs; (d) I am willing to stand up for whatever I find right; (e) Generally, I try to translate my principles into action.

The second scale is *Stability of Attitudes*. A high score on *Stability* means that the individual is willing to maintain his or her attitudes across time and believes that he or she really does so. This factor is assessed by the following items: (a) I rarely change my opinions; (b) If one has developed an attitude, one should keep it; (c) Whenever I have developed an opinion about something, I stick to it; (d) It is important for me to maintain my values.

The third scale is *Robustness against Situation Effects*. Subjects with high scores are likely to claim that their behaviour depends more on their attitudes than on the specific situation. The following items are representative of this dimension: (a) I believe that one rarely has to act contrary to one's principles; (b) Circumstances usually do not prevent me from acting in line with my attitudes.

These three facets of self-reported consistency can be reliably assessed with internal consistency coefficients (i.e., alpha) for the corresponding scales being .80 or higher.

Validation of self-report consistency scales

Validity of these three scales was investigated in two studies concerning helping behaviour and social responsibility. Details concerning the design of the studies and the measurement instruments used are provided elsewhere (Schmitt, 1988).

Study 1 was a two-year longitudinal investigation of the antecedents and consequences of the helping behaviour of adult daughters towards their mothers. A total of 807 daughters and 96 mothers took part as subjects in this project (c.f., Schmitt, Dalbert, & Montada, 1986).

A number of variables measured in this study were used in the present study: (a) *Attitudes* of a daughter towards daughters helping their mothers (in general); (b) *Personal Norms* of a daughter to provide help to her mother when she needs it (specifically); (c) *Anticipated Costs* of helping from the daughter's perspective; (d) *Intention* to help; (e) *Behaviour*—Amount of help provided by daughter according to daughters' and mothers' estimates; and, (f) *Emotional Reactions* of a daughter to her behaviour (pride, shame, guilt, etc.). These variables were related to 34 specific needs or desires a mother might address to her adult daughter (e.g., the need for contact, the need for emotional support during a crisis, or the need for care in case of illness). Each daughter was asked to select from these 34 needs those five that were most important at present. Besides these need-related variables, denial of adult children's responsibility for their parents was assessed as a general defensive tendency.

Study 2 was conducted to examine explanatory factors accounting for individual differences in social responsibility felt by people for disadvantaged members in society (physically handicapped persons, Turkish guest workers living in the Federal Republic of Germany, people living in the developing countries). The sample in this cross-sectional study included a total of 340 subjects (c.f., Montada, Schmitt, & Dalbert, 1986).

The following variables from this study were used for the present investigation: (a) *Moral Development* was measured by Lind's (1978) Moral Judgment Test; (b) *Attitudes* towards the disadvantaged encompassing emotional and behavioural components; and, (c) *Denial of Responsibility* for the disadvantaged. In addition, four justice scales were administered to subjects—preference for the *Equity Principle*, the *Need Principle*, the *Equality Principle*, and *Belief in a Just World*. Furthermore, subjects were asked which *Political Party* they had voted for in the elections for the German Parliament in 1980 and 1983. The expected relevance of this latter information was based on the assumption that preferences for certain political parties would covary with reactions to disadvantaged people, which actually proved to be the case.

Variables whose actual consistency provide theoretically meaningful criteria for determining the validity of self-reported consistency were

selected from these two studies. A few examples are offered to illustrate the procedure for assessing validity.

The validity of the first self-report scale, *Centrality*, was tested against: (a) The consistency of attitudes or personal norms and the intention to help or the amount of help actually provided; (b) The internal consistencies of various homogeneous scales (measuring attitudes, personal norms, responsibility denial, preferences for justice principles, belief in a just world, etc.); and, (c) The consistency between the emotional and behavioural components of attitudes towards the disadvantaged.

The validity of the second self-report scale, *Stability*, was tested against: (a) The actual stability of attitudes, personal norms, and denial of responsibility across 9 and 18 month time periods; and, (b) The actual stability of the political party voted for in the 1980 and 1983 elections of the German Parliament.

The validity of the third self-report scale, *Robustness*, was tested against the consistency between anticipated costs of helping one's mother on the one hand and the intention to help or the amount of help actually provided on the other hand. Because individuals with a high score on *Robustness* claim to act in line with their attitudes even under unfavourable circumstances, a negative relation was expected between *Robustness* and the two actual consistencies examined.

When it was possible, validity checks were performed using two different techniques.

Interindividual definition of consistency: In this type of analysis, self-reported consistency was treated as a variable moderating the degree of consistency between, for example, individuals' attitudes and behaviour. In doing so, multiple regression analyses with moderator variables were performed to estimate the sizes of their effects on predictor-criterion relationships (c.f., Cohen, 1978; Cronbach, 1987; Saunders, 1956). For example, to test the validity of self-reported *Attitudinal Stability*, attitudes at time 2 were regressed on attitudes at time 1, self-reported *Stability of Attitudes*, and the interaction effect involving the product of attitudes at time 1 by self-reported *Stability*. The partial correlation between the product (i.e., predictor by moderator) and the criterion variable is the moderator effect of interest in such analyses (c.f., Cohen, 1978; Dalbert & Schmitt, 1986).

Intraindividual definition of consistency: Following Magnusson and Endler (1977), we also examined an intraindividual definition of consistency or coherence. In this type of analysis, actual consistencies were determined within individuals for the pairs of variables mentioned above as well as others that were theoretically meaningful. For these analyses, four different coefficients were obtained to indicate how similar two response profiles were in terms of elevation,

scatter, and shape, as well as in a global sense (c.f., Schmitt, 1989). For example, similarity in elevation was measured by the absolute mean difference between the two variables compared, whereas the well known geometrical distance D was used as a measure of global similarity. Then, the validity of the self-report measures of consistency was tested by regressing these similarity coefficients on the former. For example, the similarity in shape of daughters' Time 1 and Time 2 attitude profiles across the 34 potential needs of their mothers was correlated with daughters' self-reported *Stability of Attitudes*.

In all analyses, possible moderator effects of *Social Desirability and Acquiescence* were controlled for statistically because both were correlated with the self-report measures of consistency and because both have been found to moderate consistency. In regression analyses with moderator variables, the effects of these variables were controlled by including them as predictors as well as their products with the moderator (self-report consistency)--in addition to the predictor and the specific product at issue theoretically. In the second type of analyses, the intraindividual profile similarities were not only regressed on self-reported consistency but on social desirability and acquiescence as well.

Results

Results of these analyses can be summarized by a brief and straightforward statement: Significant relations between self-reported and actual consistency did not exceed what would be expected by chance. Indeed, the strongest relation between self-reported and actual consistency amounted to only 4% of shared variance.

For example, none of the partial correlations between *Centrality* (moderator) and the various intraindividual profile similarities between attitudes or personal norms to help one's mother (predictors) and intention to help or amount of help actually provided (criteria) was larger than .09 (controlled for social desirability and acquiescence). Moreover, when the corresponding differential (interindividual) consistencies were estimated via multiple regression analyses using moderator variables, none of the four partial products (one moderator, two predictors, two criteria) was statistically significant.

A second example pertaining to *Centrality* involved examining the alpha reliabilities (i.e., internal consistency reliabilities) of various scales for groups differing in self-reported *Centrality*. In both

studies, the alpha coefficients for all homogeneous scales mentioned above (e.g., preference for the equity principle) were computed separately for subjects high and low in *Centrality* (median split). Subjects scoring high on self-reported *Centrality* were significantly more internally consistent in their responses on these scales than subjects scoring low in *Centrality* for only three of 22 comparisons.

Results for self-reported *Stability of Attitudes* were equally disappointing. In Study 1, assessments of attitudes to help one's mother, personal norms about helping one's mother, and denial of responsibility for the elderly were administered three times with a lag of nine months between two adjacent times of measurement. The partialled product of moderator and predictor were significant in only two of the 18 multiple regression analyses that were conducted with *Stability of Attitudes* as a moderator variable (three times of measurement by three constructs by two homogeneous subscales per construct). A very similar result was obtained when *Stability* was correlated with the actual intraindividual stabilities (i.e., profile similarities across the items of a scale) of the constructs mentioned.

In addition, subjects scoring high on *Stability of Attitudes* proved no more stable in their voting behaviour compared to subjects scoring low on self-reported *Stability*. On the basis of the political situation during the elections compared, an index was developed to measure the similarity in voting at the two elections. A value of zero was assigned to subjects who voted for the same party at each election; a value of 1 was given for moderate changes in voting behaviour--for example from a conservative (Christian Democrats) to a liberal party (Free Democrats); and, a value of 2 was used to code extreme changes--for example switching from the Christian Democrats to leftist parties such as the Social Democrats or the Greens (environmentalists). Controlling for social desirability and acquiescence, the partial correlation between this index and self-reported *Stability of Attitudes* was only -.04.

Finally, results for self-reported *Robustness* against Situation Effects showed that it did not moderate the effects of anticipated costs of helping on the intention to help and the amount of help actually provided. In the two corresponding multiple regression analyses conducted, the critical interaction terms were not significant. Moreover, the intraindividual profile similarities of costs vis-a-vis the two criteria intention and help were small in magnitude (<.09) and not significant.

Discussion and conclusions

What might account for this pattern of results - the failure of self-reported consistency to correlate with actual consistency of attitudes and behaviour as well as actual consistencies among other theoretically meaningful measures? At least in these studies, it is quite unlikely that the failure is a result of statistical or methodological problems. On the one hand, unreliability of measurement is not presumable because almost all of the measures used to determine actual consistency had estimated reliabilities of .80 or above. In addition, a lack of statistical power cannot explain the small number of significant relations between self-reported and actual consistency. Most of the critical tests had a power of above .95, so that even if effect sizes were small there was sufficient power to identify them.

Two other interpretations could be drawn from these results. First, large intraindividual differences in consistency may exist so that strong correlations between global self-reports of consistency and actual consistencies within specific attitude and value domains should not be expected. This interpretation stands in opposition, however, with the various causal explanations of individual differences in behavioural consistency that have been advanced by several theorists (e.g., Bem, 1972; Fenigstein, Scheier, & Buss, 1975; Mills & Hogan, 1978; Snyder, 1979, 1987). Furthermore, this interpretation is incompatible with studies that have found considerable generalizability of self-reported consistency across various personality traits or behavioural domains (Borkenau, 1981; Chaplin & Goldberg, 1984; Zanna & Olson, 1982: Experiment 3).

A second interpretation of the results is that self-reports of consistency are not valid - people do not know how strongly their behaviour depends on their attitudes and how stable these are. This interpretation is supported by evidence from a study on the correspondence between self- and other-rated consistency (c.f., Schmitt, 1988; Schneider, Meißner, Montada, & Reichle, 1987). In this study, 173 subjects nominated up to three friends or relatives who were then asked to rate the target subject with respect to the three facets of consistency discussed in the present research (i.e., *Centrality, Stability, and Robustness*). Ratings by other people correlated no higher than .30 with the corresponding self-ratings on these facets.

Even though the validity interpretation seems more appropriate than the specificity explanation, both lead to the same conclusion: General self-reports of attitudinal consistency - as the ones used in this study

- are not suitable to differentiate persons in terms of how well they fit a general dispositional trait model.

Author note

The research reported in this article was funded by the VW-Foundation (Stiftung Volkswagenwerk). I would like to thank Steven Cornelius for his help in revising this article.

References

- Amelang, M., Kobelt, C. & Frasch, A. (1985). Auf der Suche nach Personen mit Eigenschaften: Untersuchungen zur Restriktion des Eigenschaftsmodells auf Untergruppen von Personen, Verhaltensweisen und Situationen. In D. Albert (Ed.), *Bericht über den 34. Kongreß der Deutschen Gesellschaft für Psychologie in Wien 1984*. Band 1 (pp. 330-333). Göttingen: Hogrefe.
- Amelang, M. & Borkenau, P. (1981). Vorhersagen über einige Personen und viele Merkmale. Oder: Konsistenz über Merkmale und Kontextbedingungen als Eigenschaft. In W. Michaelis (Ed.), *Bericht über den 32. Kongreß der Deutschen Gesellschaft für Psychologie in Zürich 1980* (pp. 495-498). Göttingen: Hogrefe.
- Asendorpf, J. (1988). *Die differentielle Sichtweise in der Psychologie - angewandt auf die Entwicklung sozialer Gehemmtheit*. Unpublished Habilitation, Universität München.
- Bem, D.J. (1972). Constructing cross-situational consistencies in behaviour: Some thoughts on Alker's critique of Mischel. *Journal of Personality*, 40, 17-26.
- Bem, D.J. & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistency in behaviour. *Psychological Review*, 81, 506-520.
- Borkenau, P. (1981). *Intraindividuelle Variabilität und differentielle Vorhersagbarkeit*. Unpublished Dissertation, Universität Heidelberg.
- Chaplin, W.F. & Goldberg, L.R. (1984). A failure to replicate the BEM and ALLEN-study of individual differences incross-situational consistency. *Journal of Personality and Social Psychology*, 47, 1074-1090.
- Cohen, J. (1978). Partialled products are interactions; partialled powers are curve components. *Psychological Bulletin*, 85, 858-866.
- Cronbach, L.J. (1987). Statistical tests for moderator variables: Flows in analyses recently proposed. *Psychological Bulletin*, 102, 414-417.

- Dalbert, C. & Schmitt, M. (1986). Einige Anmerkungen und Beispiele zur Formulierung und Prüfung von Moderatorhypothesen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 7, 29-43.
- Emmons, R.A. & Diener, E. (1986). Situation selection as a moderator of response consistency and stability. *Journal of Personality and Social Psychology*, 51, 1013-1019.
- Fazio, R.H. & Zanna, M.P. (1981). Direct experience and attitude-behaviour consistency. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*. Vol. 14 (pp. 161-202). New York: Academic Press.
- Fenigstein, A., Scheier, M.F., & Buss, A.H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, 42, 522- 527.
- Greaner, J.L. & Penner, L.A. (1982). The reliability and validity of Bem and Allen's measure of cross-situational consistency. *Social Behaviour and Personality*, 10, 241- 244.
- Lind, G. (1978). Wie mißt man moralisches Urteil? Probleme und alternative Möglichkeiten der Messung eines komplexen Konstruktes. In G. Portele (Ed.), *Sozialisation und Moral* (pp. 171-201). Weinheim: Beltz.
- Lippa, R. & Mash, M. (1981). The effects of self-monitoring and self-reported consistency on the consistency of personality judgments made by strangers and intimates. *Journal of Research in Personality*, 15, 172-181.
- Magnusson, D. & Endler, N.S. (1977). Interactional psychology: Present status and future prospects. In D. Magnusson & N.S. Endler (Eds.), *Personality at the crossroads* (pp. 3-31). Hillsdale, NJ: Erlbaum.
- Mills, C. & Hogan, R. (1978). A role theoretical interpretation of personality scale item responses. *Journal of Personality*, 46, 778-785.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Mischel, W. & Peake, P.K. (1982). Beyond déjà vu in the search for cross-situational consistency. *Psychological Review*, 89, 730-755.
- Monson, T.C., Hesley, J.W., & Chernick, L. (1982). Specifying when personality traits can and cannot predict behaviour. An alternative to abandoning the attempt to predict single act criteria. *Journal of Personality and Social Psychology*, 43, 385-399.
- Montada, L., Schmitt, M., & Dalbert, C. (1986). Thinking about justice and dealing with one's own privileges: A study of existential guilt. In H.W. Bierhoff, R. Cohen, & J. Greenberg (Eds.), *Justice in social relations* (pp. 125-143). New York: Plenum Press.

- Paunonen, S.V. & Jackson, D.N. (1985). Idiographic measurement strategies for personality and prediction: Some unredeemed promissory notes. *Psychological Review*, 92, 486-511.
- Rushton, J.P., Jackson, D.N., & Paunonen, S.V. (1981). Personality: Nomothetic or idiographic? A response to Kenrick and Stringfield. *Psychological Review*, 88, 582- 589.
- Saunders, D.R. (1956). Moderator variables in prediction. *Educational and Psychological Measurement*, 16, 209-222.
- Schmitt, M. (1988). *Über die Konsistenzannahme des Eigenschaftsmodells und über Konsistenz als Eigenschaft: Theoretische, methodische und empirische Untersuchungen*. Unpublished Dissertation, Universität Trier.
- Schmitt, M. (1989). Ipsative Konsistenz (Kohärenz) als Profilähnlichkeit. *Trierer Psychologische Berichte*, 16, Heft 2.
- Schmitt, M., Dalbert, C. & Montada, L. (1982). *Interpersonale Verantwortlichkeit erwachsener Töchter ihren Müttern gegenüber: Rekrutierung der Ausgangsstichprobe, Erhebungsinstrumente in erster Version und Untersuchungsplan* (Berichte aus der Arbeitsgruppe 'Verantwortung, Gerechtigkeit, Moral' Nr. 14). Trier: Universität Trier, Fachbereich I - Psychologie.
- Schmitt, M., Dalbert, C., & Montada, L. (1986). Prosoziale Leistungen erwachsener Töchter gegenüber ihren Müttern: Unterschiede in den Bedingungen von Absicht und Ausführung. *Psychologische Beiträge*, 28, 139-163.
- Schneider, A., Meißner, A., Montada, L. & Reichle, B. (1987). *Validierung von Selbstberichten über Fremdratings* (Berichte aus der Arbeitsgruppe 'Verantwortung, Gerechtigkeit, Moral' Nr. 41). Trier: Universität Trier, Fachbereich I - Psychologie.
- Snyder, M. (1979). Self-monitoring processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol.12, pp. 86-128). New York: Academic Press.
- Snyder, M. (1987). *Public appearances, private realities: The psychology of self-monitoring*. New York: Freeman & Co.
- Wymer, W.E. & Penner, L.A. (1985). Moderator variables and different types of predictability: Do you have a match? *Journal of Personality and Social Psychology*, 49, 1002- 1015.
- Zanna, M.P. & Olson, J.M. (1982). Individual differences in attitudinal relations. In M.P. Zanna, E.T. Higgins, & C.P. Herman (Eds.), *Consistency in social behaviour* (pp. 75-103). Hillsdale, NJ: Erlbaum.
- Zuckerman, M., Bernieri, F., Koestner, R., & Rosenthal, R. (1989). To predict some of the people some of the time: In search of moderators. *Journal of Personality and Social Psychology*, 57, 279-293.
- Zuckerman, M., Koestner, R., DeBoy, T., Garcia, T., Maresca, B.C., & Sartoris, M. (1988). To predict some of the people some of the

time: A reexamination of the moderator variable approach in personality theory. *Journal of Personality and Social Psychology*, 54, 1006-1019.

Ambiguity of Italian probability expressions: Biasing effect of context

Franca Agnoli & Federica Durano

Everyday life frequently requires informal reasoning about probabilistic events even though formal education in probabilistic reasoning is considered a difficult advanced topic. The large number of probabilistic expressions in common usage (such as rarely, probably, seldom, possibly) are evidence of the importance of informal probabilistic reasoning. To communicate and reason effectively about probabilistic events, these expressions should have mutually understood, reasonably precise interpretations. Informal interviews suggest, however, that these expressions are generally vague and imprecise.

Experiment 1

Experiment 1 was conducted to explore both the precision of probabilistic expressions and inter-subject agreement about their meanings. Subjects were asked to define the meanings of 30 Italian probabilistic expressions by assigning either a single numerical value or a range of values to each expression. These probabilistic expressions and their English translations are shown in Table 1. The subjects were 78 Italians between 18 and 24 years of age. They were instructed to define each probabilistic expression by marking a point or an interval on a graduated line spanning 0 (expressing certainty that an event would not occur) to 100 (expressing certainty that an event would occur).

Figures 1a and 1b show the median and interquartile range (25th to 75th percentile) for each expression. These figures were constructed by replacing each reported interval by its midpoint, then analyzing these midpoints and the reported single values. A striking feature of these results is the wide range of probabilities represented by some of