Int. J. of Human Resource Management 15:3 May 2004 475–492

The Global Leadership Life Inventory: development and psychometric properties of a 360-degree feedback instrument

Manfred F.R. Kets de Vries, Pierre Vrignaud and Elizabeth Florent-Treacy

Abstract The purpose of this paper is to describe the design of Global Leadership Life Inventory (GlobeInvent), a 360-degree leadership feedback instrument. This instrument is presently used in executive programmes to help identify the operational mode of individual executives. Proper use of this instrument enables the user to determinate those areas of leadership behaviour where improvement is needed.

Because most studies pay attention only to the surface manifestations of leadership, most leadership feedback instruments, in turn, are not concerned with the psychodynamic processes that underlie leaders' character and behaviour. To address this gap, the GlobeInvent is based on a clinical orientation to the study of leadership. This approach provides a more complete analysis of the 'inner theatre' of leaders – that is, what makes them tick – as well as measuring the dynamic, two-way relationship between leaders and followers.

The first step in designing the instrument was to pinpoint significant themes pertaining to exemplary leadership. To that end, semi-structured interviews with senior executives were held. The leadership dimensions that emerged from that process were then tested on an international sample of senior executives. Analysis of the data from the testing confirmed the existence of twelve robust dimensions with a high reliability and internal consistency.

Because the GlobeInvent is a 360-degree feedback instrument, this article addresses differences between 'Self' scores and scores given by others ('Observers'), gender differences in scoring and the influence of nationality, management experience and age on test results. The implications of using such an instrument as a 360-degree feedback tool are reviewed, and suggestions for future research are offered.

Keywords 360-degree leadership questionnaire; global leadership; developing leadership; measuring leadership effectiveness; dimensions of leadership; GlobeInvent.

Leadership: a definitional confusion

The Anglo-Saxon etymological origin of the words *lead*, *leader* and *leadership* is *læd*, which means 'path' or 'road'. The verb *læden* means 'to travel'. Thus a *leader* is one who shows fellow travellers the way by walking ahead. This age-old metaphor of the leader as helmsman is still applicable. Unfortunately, the clarity of the etymology of the word

Manfred F.R. Kets de Vries, Raoul de Vitry d'Avaucourt Clinical Professor of Leadership Development, INSEAD, France & Singapore. Pierre Vrignaud, Doctor of Psychology, Researcher at l'INETOP/CNAM, France. Elizabeth Florent-Treacy, Research Project Manager, INSEAD, France & Singapore.

The International Journal of Human Resource Management
ISSN 0958-5192 print/ISSN 1466-4399 online © 2004 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0958519042000181214

leadership far exceeds the clarity of the concept itself. Papers, books and articles claiming to identify the ingredients of effective leadership proliferate, yet their conclusions about what leaders do are often confusing and even conflicting. Indeed, one of the major scholars of leadership has observed that 'there are almost as many definitions of leadership as there are persons who have attempted to define the concept' (Bass, 1990). Stogdill (1974), in his *Handbook of Leadership*, reviewed seventy-two definitions proposed by leadership researchers between 1902 and 1967. Among the more widely accepted of these were descriptions of what makes for effective leadership in terms of traits, behaviour, relationships and follower perceptions.

Too many of the myriad leadership studies focus on social phenomena other than their original subject of investigation. Rather than concentrating on what key decision-makers at the strategic apex of their organization are doing in the context of their work environment, researchers all too frequently draw their major conclusions from laboratory experiments, observations of leaderless groups or the activities of lower-level supervisors. *If leadership is to be a viable area of study – and if that study is to be of service to a constituency of executives – its research focus needs to be closely tied to observations of the behaviour and actions of individuals in leadership positions*. The impetus to develop the *Global Leadership Life Inventory* (*GlobeInvent*) originated in the desire to understand what exemplary leaders actually *do*.

Because most studies pay attention only to the surface manifestations of leadership, most leadership feedback instruments, in turn, are not concerned with the psychodynamic processes that underlie leadership 'rationale'. To address this gap, the GlobeInvent is based on a *clinical* orientation to the study of leadership. This approach provides a more complete analysis of the 'inner theatre' of leaders – that is, what makes them tick – as well as measuring the dynamic, two-way relationship between leaders and followers.

Furthermore, most existing leadership assessment instruments are designed only for *self*-assessment (an inaccurate process, since respondents tend to colour their responses according to how they would like to be perceived – the so-called social desirability factor); they do not incorporate assessment by others. Research clearly indicates that 360-degree feedback systems give a much more accurate picture than self-assessment of what executives really do and how executives actually behave (London *et al.*, 1990; Hazucha *et al.*, 1993; Kluger and DeNisi, 1996; Walker and Smither, 1999). However, there are very few reliable and valid *360-degree* instruments in existence (Van Velsor and Leslie, 1991a, 1991b). Because there is frequently a significant difference between what leaders *say* that they do and what they *actually* do, 360-degree feedback became the operational method of choice for the GlobeInvent.

A clinical orientation in the design of a global leadership instrument

Broadly speaking, two extreme positions can be identified among scholars of leadership. On one side of the spectrum are the 'personalists' – researchers who argue that specific personality variables determine leadership effectiveness. On the other side of the spectrum are the 'situationists' – those who deny the influence of individual differences and attribute all variations in leadership effectiveness to environmental constraints. While personalists view leaders as heroic helmsmen, in control of whatever situation they find themselves in, situationists turn leaders into figureheads – puppets manipulated by forces of the environment. Those in the former camp rarely dig below the surface to analyse *why* a particular helmsman becomes a hero. Those in the latter camp claim that it makes little difference what a leader does; societal forces, they say, determine what actions should be (and are) taken.

These opposing positions (and their melding in an interactionist perspective) set the stage for a cornucopia of theories, each backed by strong defenders. One can find 'great man' theories, trait theories, situational theories, psychoanalytic theories, political theories, humanistic theories, cognitive theories, leader-role theories, reinforced change theories, path-goal theories, contingency theories, multiple linkage theories, vertical dyad linkage theories, exchange theories, behavioural theories and attribution theories. (It is unnecessary – and would indeed be impossible – to recapitulate all these theories here. For a thorough overview, see Bass (1990), Yukl (1994) and House and Aditya (1997).)

Although new research into leadership has a widening scope, most studies still pay attention primarily to the visible manifestations of human action and/or societal influences (Posner and Kouzes, 1988; Bass and Avolio, 1990a, 1990b; Posner and Kouzes, 1993; Kouzes and Posner, 1995; Alimo-Metcalfe and Alban-Metcalfe, 2001). Little effort has been made to enrich these approaches to leadership with a more clinical orientation.

The *clinical orientation* to leadership uses findings from psychoanalysis, cognitive theory, developmental psychology and family systems theory to arrive at a more complete understanding of the dynamic process that exists between leaders and followers. Advocates of the clinical orientation to leadership argue that deconstructing the major preoccupations of executives – what is sometimes called their 'inner theatre' – helps illuminate the major themes that drive behaviour. In decoding their human 'texts', researchers extract significance from interrelated behavioural, cognitive and affective manifestations that have developed out of an individual's experiences. In the deconstruction of the dynamics of leadership, this orientation looks to the *triangle of mental life*, consisting of emotion, cognition, and behaviour. While other approaches to leadership focus on the latter two elements, the clinical approach includes emotions in the equation.

The clinical approach puts a broad relational slant on the study of leadership. With the evolution of the clinical approach, scholars of psychoanalysis and leadership have deepened the understanding of the complex, interactive role of the leader in society (Erikson, 1958, 1969; Zaleznik and Kets de Vries, 1975; Kohut, 1977; Kohut, 1985; Zaleznik, 1977; Kets de Vries, 1993, 2001).

As studies in the clinical approach to leadership have deepened the insight into the behaviour of executives, a new wind has blown in with contributions from leadership researchers influenced by the findings of political science. Their way of looking at leadership centres on a renewed interest in the concept of *charisma* – that 'divinely inspired gift' attributed to leaders by their followers. With the leader increasingly seen as a crucial agent of change in a fast-changing world, charismatic leadership models are far more in tune with contemporary reality (and thus are more helpful) than are models offered by earlier leadership research.

The current emphasis on charisma – on the *inspirational, transformational role of leaders* – can be viewed as a paradigm shift in the study of leadership. Present-day researchers are turning to the study of leaders who *by force of their personality* have an extraordinary effect on their followers. It is becoming increasingly clear that the primary challenge for present-day leaders lies in affecting the mindset of organizational participants through value creation, through influencing the organization's culture and through building commitment to the organization's vision, mission and strategies to obtain well-above-average organizational performance (Zaleznik, 1977; Bennis and Nanus, 1985; Tichy and Devanna, 1986).

However, insightful as this new trend in research may be, most leadership feedback instruments are still not concerned with the psychodynamic processes that underlie the leadership puzzle. In addition, most are not 360° instruments, and therefore an important source of information – and essential 'reality check' – is lacking. To understand what

leaders really do, and to fill this gap in leadership assessment, the first author decided to study top executives who participated in a leadership programme at INSEAD ('The Challenge of Leadership: Developing Your Emotional Intelligence').

Leaders who are selected to participate in this seminar are typically at the top of their career ladder. They come to the programme with the intention of taking some time to reflect on their life goals, both professional and personal. Therefore, this group of leaders was the ideal place to begin testing our theories on leadership behaviour, and to distil the theories into measurable dimensions for a 360° instrument. Over the course of a large number of iterations of this programme, more than 300 senior executives were asked what issues were most important to them in their day-to-day work, and what kind of behaviour contributed to their effectiveness. They were also asked to discuss concerns about their personal life. In addition, their own responses were compared with the perceptions of others (i.e. peers, subordinates, superiors, customers, spouses, children, other family members and friends). From the feedback it became clear that congruency existed between self-perception and the perception of others in only a minority of the cases (Dalessio, 1998). The observation of these variances served as a catalyst for the development of a 360-degree feedback instrument that would reflect the *actual* behaviour of executives, highlighting behaviour patterns that made for leadership effectiveness.

Exploratory interviews for the development of the instrument were conducted in a semi-structured fashion over a period of three years. Each respondent was approached with a list of open-ended questions pertaining to major concerns. Depending on the responses of the group as a whole, questions were dropped, revised or retained. Supplemental observational data were collected by the principal author in the form of notes taken while studying the various executives in meetings and while participating in a large number of action research projects and strategic interventions. In the course of this fieldwork and instrument development, 'grounded theory' was used to arrive at a set of hypotheses about various preoccupations of the participating executives; in other words, while engaged in the process of hypothesis formulation, the researchers delineated connections, patterns and themes, continuously modifying their hypotheses as dictated by emerging material (Glaser and Strauss, 1967; Argyris and Schon, 1974). The observed patterns of behaviour were then integrated with knowledge about the growth and development of human beings and the findings of developmental and clinical psychologists on the functioning of human personalities. Through this ethnographic and clinical orientation, ideas were developed and 'thick' description emerged – that is, description that involved 'guessing at meanings, assessing the guesses, and drawing explanatory conclusions from the better guesses', to use Clifford Geertz's words (1973: 20). In addition, perceptual distortions due to participant observation were explored (Devereux, 1978; Van Maanen, 1988; Schein, 1987).

Method: operationalizing the pilot instrument

Sample

The total sample used for the construction of the *Global Leadership Life Inventory* consisted of three groups of executives representing countries from all over the world:

1. *CEO sample* Participants in the INSEAD top management seminar 'The Challenge of Leadership: Developing Your Emotional Intelligence' participated in the research project. These participants were also interviewed to assess their perceptions of what

made for effective leadership. In addition, two faculty members associated with the programme made a separate assessment of the participants' effectiveness. Fifty CEOs completed the questionnaire.

2. *INSEAD MBA sample* One hundred and fifty MBA candidates completed the questionnaire.
3. *INSEAD's Advanced Management Programme sample* One hundred and twenty senior executives participating in the Advanced Management Programme completed the questionnaire. A considerable number of these executives also asked colleagues, customers, family members and others to assess them via the questionnaire.

Analysis of the constructs

The constructs that emerged from the in-depth interviews with top executives were content-analysed by five faculty members working independently of each other, and then grouped in terms of themes relevant to leadership. Inspection of the groupings formed by the researchers revealed a very high level of agreement among faculty members. The constructs of the inventory were derived through triangulation of the data in the group discussion.

These interviews confirmed our belief that the most effective global leaders play two roles simultaneously: the first is a charismatic role; the second is what we might call 'architectural'. The first involves envisioning, empowering and energizing – behaviours that direct, inspire and motivate their followers. The second involves the implementation of processes to improve organizational design and to control and reward employee behaviour appropriately (Kets de Vries and Florent-Treacy, 1999). The most successful leaders appear to be extremely talented at aligning their charismatic with their architectural role.

These two roles – never easy to balance effectively – are all the more difficult to accomplish successfully in the context of a global organization. In addition, we found that leaders of global organizations had specific skills that enabled them to adapt their architectural and charismatic roles to different constituencies. They also were able to balance the requirements of the different stakeholders outside their organization. They had a set of personal characteristics, such as hardiness and tenacity, that made them tougher than most. We determined that they were able to understand and govern their own emotions, as well as those of their followers. Finally, we recognized the importance of life balance to the success of top leaders – as Freud put it succinctly, they focused on *work* and *love* in balanced measure, giving adequate attention to the requirements of their personal lives.

In fact, we determined, with the help of the triangulation process, that executives of world-class organizations focus on twelve main tasks conceptualized here as dimensions:

1. Articulating a compelling vision, mission and strategy with a multi-country, multi-environment, multi-function and multi-gender perspective that connects employees, shareholders, suppliers and customers on a global scale.
2. Giving workers at all levels a voice by empowering them through the sharing of information and the delegation of decisions to the people most competent to execute them.
3. Energizing and motivating employees to actualize the organization's specific vision of the future.
4. Creating the proper organizational design and control systems to make the guiding vision a reality, and using those systems to align the behaviour of the employees with the organization's values and goals.

5. Setting up the appropriate reward structures and giving constructive feedback to encourage the kind of behaviour that is expected from employees.
6. Creating team players and focusing on team effectiveness by instilling a co-operative atmosphere, building collaborative interaction and encouraging constructive conflict.
7. Making employees aware of their outside constituencies, emphasizing particularly the need to respond to the requirements of customers, suppliers, shareholders and other interest groups, such as local communities affected by the organization.
8. Inculcating a global mentality in the ranks – that is, instilling values that act as a sort of glue between the regional and/or national cultures represented in the organization.
9. Encouraging tenacity and courage in employees by setting a personal example in taking reasonable risks.
10. Fostering trust in the organization by creating, primarily through example, an emotionally intelligent workforce whose members know themselves and know how to deal respectfully and understandingly with others.
11. Articulating and modelling the importance of the need for life balance for the long-term welfare of employees.
12. Paying attention to work, career, life and health stress issues, and balancing appropriately the various kinds of pressures that life brings.

These themes were made operational by constructing clusters of questions that were given the following descriptions: *Envisioning, empowering, energizing, designing and controlling, rewarding and giving feedback, team-building, outside stakeholder orientation, global mindset, tenacity, emotional intelligence, life balance and resilience to stress*.

Designing questionnaire items

Working initially independently of each other, five faculty members devised a series of statements that reflected the constituent constructs in each cluster. These statements again were triangulated during group discussions. Emphasis was given to behavioural dimensions and specificity of each statement.

Developing the pilot questionnaire

In order to turn the statements about leadership into a pilot questionnaire with a minimal amount of ambiguity about each of the items, the questionnaire was pre-tested. This pre-test was particularly important, because for many of the respondents English was not the first language (and thus clarity was essential). Each item presented readily observable or inferable behaviour and was presented positively. The initial pilot questionnaire was distributed to twenty CEOs, who were asked to comment on the clarity and the face validity of the questions. After two pre-tests (the second with another sample of CEOs) the number of questionnaire items was narrowed down to 109 items. (See Appendix A for a full description of each dimension.)

360-degree feedback

The fact that self-report evidence is frequently influenced by the social desirability factor was taken into account during the design and development of the questionnaire. By incorporating 360-degree feedback, the *Global Leadership Life Inventory* capitalizes on multiple perspectives, confirming assumptions about strengths and helping identify areas that need improvement. Test-takers include in the feedback circle not only their

own self-assessment but also observations made by direct reports, peers, internal or external customers, superiors and family members, who work with an 'Observer' version of the GlobeInvent questionnaire. Participating respondents first complete the GlobeInvent-Self form and then ask a number of people familiar with their behaviour to complete the GlobeInvent-Observer form.

Biographical and organizational data

The respondents are also asked to provide the following biographical information: gender, age, nationality and professional background (current position, industry).

Validation study

The reliability of the *Global Leadership Life Inventory* was studied in three stages: 1) by checking the internal consistency within each dimension using a classic psychometric approach; 2) by evaluating the structure of the questionnaire through confirmatory factor analysis at both the item level and the dimension level; and 3) by conducting a study of the effects of the characteristics of subject groups (gender, nationality, age and experience).

The questionnaire included 109 items spread over twelve dimensions. (The number of items in each dimension is indicated in column 2 of Table 2.) Each item is presented in the form of a statement. Test-takers are asked simply to indicate (on a seven-point Likert-type scale) the degree to which each statement describes the way they act in a particular situation. As a guideline, they are advised that a score of 1 means that the statement does not describe them at all; in other words, they *never* act in the way described. A score of 7 means that they *always* act in the way described; in other words, they are *exemplary* in this area. (See Table 1 for selected questions taken from the GlobeInvent.) They are also advised that a score of 4 corresponds to 'average' behaviour for top executives.

The sample used for the validation study of this survey instrument was constructed using data from six groups of executives, from a wide range of nationalities and cultures, who attended executive training programmes on the INSEAD campus in 2001 and completed the questionnaire. Over half these 320 test-takers were outstanding executives on a global level, many of them senior executives and CEOs. A small part of the test-takers were students enrolled in INSEAD's MBA programme, and as they typically had less experience than the other respondents this population will be used to test hypotheses about the effect of age and experience in the criterion-related validation study.

All the above test-takers completed the GlobeInvent-Self form. In addition, forty-eight of them also asked an average of five or six observers to complete the GlobeInvent-Observer form on their behalf, testing the 360-degree feedback component of the instrument. The observers were classified within the following categories: peers, subordinates, superiors, partners, others. The validity study is therefore based on the analysis of 607 questionnaires (320 GlobeInvent-Self and 287 GlobeInvent-Observer).

Results

Internal consistency analysis

The means and standard deviations appear in columns 4 and 5 of Table 2. Because the number of items is different for each scale, column 3 shows the theoretical maximum score per scale. The average scores are fairly high in general, indicating a ceiling effect due to the frequency of responses on the right side of the scale. This tendency towards higher scores is a result partly of the social desirability factor and partly of the fact that

Table 1 *Selected questions from the Global Leadership Life Inventory, paired with their dimensions*

Q 8	I inspire my people to look beyond existing boundaries.	(Envisioning)
Q 10	I always try to involve my employees in decision-making.	(Empowering)
Q 22	I mobilize people to get things done.	(Energizing)
Q 25	I set clear performance standards and goals.	(Designing and controlling)
Q 38	I make sure that achievements are recognized along the way.	(Rewarding and feedback)
Q 47	I make sure that all participants feel that they contribute to to the decision-making process.	(Team-building)
Q 52	I make sure that customer satisfaction stands central.	(Outside orientation)
Q 63	I am good at adapting to business practices in cultures other than my own.	(Global mindset)
Q 65	I am prepared to stick to an unpopular decision if I feel that it is the right one.	(Tenacity)
Q 79	I work to generate trust among my people.	(Emotional intelligence)
Q 89	I set priorities in both my private and my professional life.	(Life balance)
Q 92	I feel a lot of pressure at work.	(Resilience to stress)

Table 2 *Means, standard deviations, and reliability indexes for the scales of the GlobeInvent estimated on the whole sample* ($n = 607$)

Scale	*Number of Items*	*Maximum Score*	*Mean*	*Standard deviation*	*Cronbach's alpha*
Envisioning	8	56	44.52	5.61	.77
Empowering	8	56	42.64	6.08	.80
Energizing	9	63	48.40	7.07	.82
Designing	7	49	35.26	6.30	.84
Rewarding	8	56	41.45	7.16	.87
Team-building	13	91	67.12	10.32	.85
Outside orientation	6	42	33.28	4.98	.82
Global mindset	8	56	43.82	7.98	.87
Tenacity	5	35	28.89	3.96	.76
Emotional intelligence	17	119	89.26	13.28	.91
Life balance	9	63	47.92	7.97	.79
Resilience to stress	11	77	34.77	11.44	.84

the majority of the subjects in the database were extremely high achievers (many of them CEOs of global companies, as noted earlier). These high-score issues will be explored and discussed later.

Internal reliabilities, assessed through standardized Cronbach's alpha (see column 6, Table 2), range from .76 to .91 for the different dimensions of the GlobeInvent. These values are all higher than the .70 value generally considered to indicate a sufficient reliability by classical psychometric treatises (for example, Nunnally, 1978) and by standard practice within the scientific community (on this point, see the review by Peterson, 1994).

The differences between coefficient values are partially due to the differences in number of items per scale. In fact, research has found that the number of items in the calculation of alpha coefficients can appear to create confusion between internal consistency and the length of the scale (Cortina, 1993). In order to compare the reliability of scales of different lengths, we recalculated the alpha coefficients using the same number of items for each scale and the mean correlation between items of the original scale. The internal reliability for the scales containing ten items ranges from .80 to .90. These values show that the weaker alphas for certain scales are directly related to the smaller number of items in those scales; they are not a result of a lower internal consistency compared to the other scales.

The internal reliability of 360-degree feedback instruments is, in general, lower for questionnaires filled out by the subjects themselves than for questionnaires filled out by their 'observers' (Posner and Kouzes, 1988, 1993; Kouzes and Posner, 1995). It is believed that the difference in internal consistency can be explained by the manifestation of the social desirability factor. For the GlobeInvent, reliability ranges from .72 to .87 for the 'Self' scores (column 2, Table 3), and from .79 to .93 for the 'Observer' scores (column 3, Table 3). Thus the observation of outsiders appears to be more reliable than self-evaluation. This phenomenon, undoubtedly related to a more reliable use of the response scale by observers, underscores the importance of using 360-degree feedback instruments.

Cronbach's alpha gives information on the reliability of the scale as a whole. It is equally important, however, to look at internal consistency at the item level. The strength of the relationship of each item to its scale can be measured by examining the corrected item test correlation (CITC). In the GlobeInvent, the CITC ranges from .26 to .77, with a median at .54. For 102 items out of the 109, the CITC is higher than .40, a value usually considered to indicate a reliable relationship between the item and the scale to which it belongs. Based on this analysis, we can conclude that 94 per cent of items are well placed in their scale.

Confirmatory factor analysis

At the item level In order to test the hypothesis that the relationship among the items can be explained using twelve factors, the structure of the questionnaire was studied using structural equation modelling. The model tested was made up of twelve factors that each corresponded to one of the twelve distinct scales. The items load on one factor only; no cross-loading was postulated. The test demonstrated a correlation between the factors and the scales.

It has often been noted that models of measurement cannot fit perfectly any data using only latent variables to explain the organization of relationships among items. The best explanation for this phenomenon is that the subjects' responses are tied not only to the latent variable measured in the questionnaire, but also to 'parasitical' issues

Table 3 *Comparisons of reliability indexes and means based on GlobeInvent 'Self' and 'Observer'*

Scale	*Self*[†] *Cronbach's alpha*	*Observer*[‡] *Cronbach's alpha*	*Self*[†] *mean (SD)*	*Observer*[‡] *mean (SD)*	*Difference mean*[a]	*d*[b]
Envisioning	.73	.79	45.25 (4.77)	43.71 (6.34)	1.53***	0.27
Empowering	.72	.82	43.98 (4.88)	41.14 (6.89)	2.84***	0.47
Energizing	.76	.84	49.51 (5.69)	47.16 (8.18)	2.34***	0.33
Designing	.81	.87	35.40 (5.79)	35.11 (6.83)	0.29	0.05
Rewarding	.82	.89	42.75 (5.90)	40.00 (8.11)	2.75***	0.38
Team-building	.84	.88	68.34 (8.80)	65.76 (11.65)	2.58***	0.25
Outside orientation	.77	.86	33.43 (4.16)	33.12 (5.76)	0.31	0.06
Global mindset	.87	.88	43.53 (7.99)	44.14 (7.96)	−0.60	−0.08
Tenacity	.73	.79	29.14 (3.56)	28.61 (4.36)	0.53	0.13
Emotional intelligence	.87	.93	89.84 (10.94)	88.61 (15.48)	1.23	0.09
Life balance	.78	.80	48.02 (8.24)	47.82 (7.68)	0.20	0.02
Resilience to stress	.81	.86	35.71 (10.72)	33.72 (12.13)	1.99*	0.17

Notes
[†] $n = 320$.
[‡] $n = 287$.
[a] Mean difference $= (m_{Self} - m_{Observer})$.
[b] Effect size: $d = (m_{Self} - m_{Observer})/S_{whole}$.
* Significant at .05.
*** Significant at .001.

associated with the way the questions are formulated and affected by the social desirability factor. To correct for these parasitical relationships we introduced 862 pairs of error co-variances among the items.

AMOS (Analysis of covariance MomentS, version 4.5 (Arbuckle and Wothke, 1999)) software was used to estimate the parameters of the model. Although the test of exact fit proved to be significant at .001: Cmin = 6078.45 with 4915 ddl., the test of close fit (RMSEA = .02) is inferior to the minimum of .05 required to prove an acceptable adjustment (Browne and Cudeck, 1993). This last hypothesis can be accepted at a high probability threshold: p (RMSEA $< .05) > .99$. Furthermore the ECVI value is inferior to that of the saturated model. These elements point to the conclusion that the model shows an acceptable adjustment to the data. The loadings of the items on their factor are for the most part superior to .40 (102 items out of the 109: 94 per cent). The theoretical structure therefore appears to reflect the organization of the data accurately. In general, the correlations among factors show significant links among the dimensions measured. These correlations are particularly high among certain groups of dimensions, which indicates the existence of second-order factors.

At the dimension level To study the organization of relationships among the dimensions, an exploratory factor analysis (principal factor analysis) was done on the dimension scores using LISREL 8.5 software (Jöreskog and Sörbom, 2000). The chi-square adjustment test shows that the quality of the adjustment does not significantly increase after the fourth factor. Varimax rotations were applied to the four retained factors (see Table 4). Our testing indicates that 'Emotional Intelligence' and 'Life Balance' load mainly on the first factor;

Table 4 *Loadings of the twelve dimensions on the four principal factors after Varimax rotation*

	Factor 1	*Factor 2*	*Factor 3*	*Factor 4*
Envisioning	0.14	0.49	0.36	0.52
Empowering	0.18	0.74	0.05	0.35
Energizing	0.26	0.64	0.23	0.34
Designing	0.10	0.64	0.41	0.09
Rewarding	0.34	0.71	0.26	0.04
Team-building	0.39	0.65	0.24	0.25
Outside orientation	0.13	0.38	0.51	0.14
Global mindset	0.29	0.23	0.20	0.43
Tenacity	0.18	0.15	0.55	0.31
Emotional intelligence	0.80	0.45	0.11	0.25
Life balance	0.41	0.12	0.32	0.17
Resilience to stress	0.06	0.06	0.07	0.32

the dimensions of 'Empowering', 'Energizing', 'Designing', 'Rewarding' and 'Team-building' on the second; 'Outside orientation' and 'Tenacity' on the third; and, finally, 'Envisioning' and 'Global mindset' on the fourth. It should be noted that the dimension of 'Resilience to stress' does not appear to have high loading on any of these four factors, which proves its relative independence from the other dimensions.

Interpretation of different characteristics of the subjects

Comparison of 'Self' and 'Observer' answers Research has confirmed that the use of a 360-degree feedback instrument often results in differences between the mean of the test-taker ('Self') and the means of his or her observers (with the former usually higher). As noted earlier, most authors explain such differences by the social desirability factor, which makes people more reality-bound in observing others than in assessing themselves. In Table 3 the means and standard deviations for both positions ('Self' and 'Observers', respectively, in columns 4 and 5) are presented. The differences in the averages (column 6, Table 3) vary from -0.60 to 2.84. Because the number of items within different dimensions varies, it is difficult to interpret the significance of the differences supported by these figures. To address this, the effect size was evaluated by calculating Cohen's 'd' (the ratio between the difference in the mean and the standard deviation of the reference population; see column 7, Table 3). Based on the levels generally adopted to judge the size of the effect (Cohen, 1992; Corroyer and Rouanet, 1994), the differences on the 'Envisioning', 'Empowering', 'Energizing', 'Rewarding and Giving feedback' and 'Team-building' scales indicate a medium effect (all these differences are significant at .01), while the difference on 'Resilience to stress' indicates a low effect (significant at .05).

Certain dimensions are more sensitive than others to differences in judgement between a subject and his or her observers. One could hypothesize that for the dimensions in which the confrontation between the subject and the members of his or her entourage is greatest, there is by implication a divergence of perspective about the actions and behaviour of the subject. The fact that all the dimensions of the GlobeInvent are not equally sensitive to this phenomenon confirms the reliability of the items and scales within the context of a 360-degree feedback approach. On the other hand, the comparison of 'Observer' averages according to their categories – subordinate ($n = 52$), peer ($n = 78$), superior ($n = 33$) – do not show a significant difference.

Table 5 *Comparisons of means based on GlobeInvent women and men*

Scale	*Women*[†] *mean (SD)*	*Men*[‡] *mean (SD)*	*Difference mean*[a]	*d*[b]
Envisioning	44.91 (5.93)	44.36 (5.50)	0.56	0.10
Empowering	42.90 (6.30)	42.47 (6.06)	0.43	0.11
Energizing	48.58 (7.75)	48.18 (6.88)	0.40	0.11
Designing	35.35 (7.34)	35.13 (5.88)	0.22	0.12
Rewarding	41.92 (7.97)	41.13 (6.90)	0.78	0.12
Team-building	67.99 (11.47)	66.54 (9.90)	1.45	0.08
Outside orientation	33.50 (5.52)	33.10 (4.81)	0.40	0.11
Global mindset	45.36 (7.91)	43.41 (7.96)	1.95**	0.14
Tenacity	28.60 (4.40)	28.95 (3.80)	−0.35	0.11
Emotional intelligence	91.40 (15.23)	88.31 (12.57)	3.09**	0.11
Life balance	48.84 (7.97)	47.57 (7.87)	1.27	0.12
Resilience to stress	33.31 (12.03)	35.30 (11.17)	−1.99	0.15

Notes
[†] $n = 151$.
[‡] $n = 433$.
[a] Mean difference $= (m_{women} - m_{men})$.
[b] Effect size: $d = (m_{women} - m_{men})/S_{whole}$.
** Significant at .01.

Gender effect To study the effect of gender on responses, the average 'Self' scores of men and women in the database were compared. The results are presented in Table 5. A significant difference exists (.01) for two dimensions: 'Global mindset' and 'Emotional intelligence'. However, the size of the effect measured by Cohen's 'd' statistic is low.

The study of interactions allows us to see if a gender-dependent difference exists as a function of the role of the respondent ('Self' or 'Observer'). The study of the differences between the averages in the 'Self' score for women ($n = 47$) and men ($n = 253$) indicates a significant difference (at .01) and a midsize effect for the dimension of 'Emotional intelligence'. To analyse the male/female differences in the observer position, it is necessary to focus not only on the gender of the observer, but also on the gender of the observed. The averages were calculated according to subjects' and observers' gender: male subject ($n = 37$) observed by men ($n = 142$) and by women ($n = 70$); or female subject ($n = 11$) observed by men ($n = 38$) and by women ($n = 34$).

Incidentally, we noticed that male subjects tend to select male observers (67 per cent of observers) more frequently, whereas women select observers of both sexes equally (53 per cent of their observers were male). It would be interesting to track whether this phenomenon is linked to the percentage of each gender in a subject's entourage. In other words, are there more men in the pool of potential observers when the CEO-subject is a man – or is it possible to hypothesize that there is an 'attraction' between the gender of the subject and the gender of the observer, in that men naturally tend to ask other men to fill out the questionnaire as their observers? In any case, this phenomenon merits further study.

Concerning the interaction between the gender of the subject and that of the observer, a comparison of the averages does not show a statistically significant difference. However, there appear to be certain tendencies at a descriptive level: male observers judge female subjects to be less accomplished in the dimensions of 'Envisioning' and 'Team-building' and more sensitive to 'Resilience to stress'; they also judge women to be more accomplished in the dimension of 'Rewarding and feedback'. Female observers

also rate female subjects to be less effective in the dimension of 'Envisioning' and more sensitive to stress. However, unlike male observers, female observers consider female subjects to be more effective at 'Team-building'. It seems that there is in fact an interaction between the sex of the subject and the sex of the observers on the perception of performance in these areas.

To summarize these observations, it can be said that the effect of gender on the questionnaire results appears to be weak for the dimensions taken as a whole. However, certain interaction phenomena between the sex of the observer and that of the subject – in particular, the opposing perspective on the ability of women to build teams – merits further exploration and explanation. We found no other studies that looked at the effect of gender in 360-degree feedback using an interaction approach.

Cross-cultural observations The GlobeInvent was constructed as a multicultural instrument; the validation database contains subjects of over forty different nationalities (see Table 6). It was difficult to study the possible impact of national culture of the subject for all of the nationalities included in the database, because some of the groups

Table 6 *Nationality of total GlobeInvent sample*

	Self	*Observer*	*Total*
Australian	4	1	5
Austrian	3		3
Belgian	9	7	16
Brazilian	7		7
British	38	63	101
Canadian	4	1	5
Chinese	4	2	6
Danish	8	4	12
Dutch	35	27	62
Finnish	4		4
French	36	61	97
German	21	19	40
Greek	3	4	7
Irish	5	8	13
Israeli	1	4	5
Italian	10	8	18
Japanese	7		7
Malaysian	3	5	8
New Zealander	3		3
Nigerian	6		6
Norwegian	3	2	5
Polish	4		4
Portuguese	4		4
Russian	2	2	4
Singapore	4		4
South African	11	5	16
Spanish	5	1	6
Swedish	5	10	15
Swiss	9	7	16
USA	8	25	33
Zambian		3	3
Total	266	269	535

were too small to allow valid comparison. Therefore, the cross-cultural analysis focused on the two largest groups, the British ($n = 101$, with 38 in the 'Self' position and 63 in the 'Observer' position) and the French ($n = 97$, with 36 in the 'Self' position and 61 in the 'Observer' position).

Concerning the 'Self' responses, significant differences appear for the dimensions of 'Empowerment' (significant at .01), 'Team-building' (at .05), 'Global mindset' (at .001) and 'Tenacity' (at .05). These differences all tend to show a higher score for the French subset. In the 'Observer' position, significant differences appear for the dimensions of 'Energizing' (at .05), indicating a higher score for the British subset. For 'Resilience to stress', however, the French 'Observer' score is higher (at .001).

The effect of experience and age The effect of cumulated experience was studied through a comparison of the results of novices and those of experienced executives. The database included a subgroup of MBA students ($n = 54$) who had an average of three years of managerial experience before starting their MBA studies. These subjects were young: thirty-eight were under 30 years old and sixteen between 30 and 34 years old. A comparison was made between the averages of this group and the averages of more experienced executives giving the 'Self' response in the database ($n = 266$); most of the latter had at least twenty years of executive experience, none of them was under 30 years old, and only thirteen of them were between 30 and 34. Significant differences (at .01) appear for the more experienced executive group in the dimensions of 'Envisioning', 'Empowering', 'Designing', 'Team-building' and 'Outside orientation', the effect ranging from medium to high (see Table 7). On the other hand, for

Table 7 *Comparisons of means based on GlobeInvent MBA candidates and executives*

Scale	*MBA candidates*[†] *mean (SD)*	*Executives*[‡] *mean (SD)*	*Difference mean*[a]	d[b]
Envisioning	43.50 (4.16)	45.60 (4.81)	−2.10***	−0.37
Empowering	42.33 (4.86)	44.31 (4.82)	−1.98***	−0.33
Energizing	49.69 (5.41)	49.47 (5.75)	0.21	0.03
Designing	32.20 (6.60)	36.04 (5.39)	−3.84***	−0.61
Rewarding	42.76 (5.81)	42.75 (5.93)	0.02	0.00
Team-building	65.64 (8.36)	68.89 (8.80)	−3.25*	−0.31
Outside orientation	32.20 (4.39)	33.68 (4.07)	−1.47*	−0.30
Global mindset	44.99 (7.88)	43.24 (7.99)	1.76	0.22
Tenacity	28.91 (3.89)	29.19 (3.50)	−0.28	−0.07
Emotional intelligence	90.11 (10.84)	89.78 (10.98)	0.33	0.02
Life balance	50.77 (7.16)	47.46 (8.34)	3.32***	0.42
Resilience to stress	32.27 (8.39)	36.41 (11.02)	−4.14**	−0.36

Notes
[†] $n = 58$.
[‡] $n = 266$.
[a] Mean difference $= (m_{MBA} - m_{Executives})$.
[b] Effect size: $d = (m_{MBA} - m_{Executives})/S_{whole}$.
* Significant at .05.
** Significant at .01.
*** Significant at .001.

the dimensions of 'Life balance' and 'Resilience to stress', significant differences appear in favour of the MBA subgroup.

The question then is: are the differences described in the preceding paragraph a result of the different level of experience for the two subgroups or a consequence of the difference in age? To resolve these alternate hypotheses, the averages for MBA candidates and executives were compared for subjects in the only age cohort where there were enough data for both groups (30–4 years old; $n = 16$ for the MBAs and $n = 13$ for the executives). There was a significant difference (at .01) in favour of the subgroup of executives in the dimensions of 'Envisioning', 'Designing', 'Team-building' and 'Outside orientation', which confirms that their better performance in these dimensions is a result of experience and not just age. It is interesting to note that there do not appear to be significant differences for the dimensions of 'Life balance' and 'Resilience to stress', which therefore seem to depend on the professional and familial responsibilities that increase as the subjects get older.

The study of the general age effect focused on the 'Self' responses of executives. Subjects were grouped into eight categories (25–9: $n = 0$; 30–4: $n = 13$; 35–9: $n = 65$; 40–4: $n = 88$; 45–9: $n = 47$; 50–4: $n = 27$; 55–9: $n = 8$; over 60: $n = 0$). The averages for the five categories that had an adequate number of subjects were compared (30–4, 35–9, 40–4, 45–9, 50–4). Because no significant difference was observed, the hypothesis of a possible age effect on any of the dimensions can be thrown out. However, with sample sizes so small, it is difficult to draw definite conclusions from this result. If the same result were to occur for a larger sample size, one conclusion that could be drawn is that the selection for candidates in these various senior executive programmes is of such rigour that only people with a high competency level are admitted, cancelling out a possible learning effect that would come with age. Because position acquired in the organization is the main criterion for admission, it appears to be a good indicator for learned competency level.

Discussion

The purpose of the study described in this study that led to the development of the *Global Leadership Life Inventory* was to investigate what it means to be a world-class leader. What kinds of behaviour and actions make for leadership effectiveness? The decision was taken to both do a literature review and adopt a qualitative methodology (in which senior executives, both male and female, described the essence of their role as a leader), thereby applying a grounded theory approach. This initial two-pronged study, through triangulation, led to the formulation of twelve dimensions viewed as essential for exemplary leaders: *Envisioning, Empowering, Energizing, Designing and controlling, Rewarding and giving feedback, Team-building, Outside orientation, Global mindset, Tenacity, Emotional intelligence, Life balance and Resilience to stress*. To make these dimensions operational for the purpose of leadership development, a quantitative research project was launched that would test the robustness of the various dimensions. Data from this study suggest that these twelve dimensions of global leadership possess high internal reliability and consistency.

Second-order factors

In general, the correlations demonstrated significant links among the various dimensions measured. That the factor scores are intercorrelated indicates that the factors have a degree of common variance. The resulting correlations are particularly high among certain groups of dimensions. When principal-components analysis was performed on

the twelve factors, a number of 'second-order' factors emerged, as noted earlier. There is a need for further analysis of the psychological significance of these factors, using factor analytic and other multivariate techniques, but these efforts are beyond the scope of this article.

Studying subgroupings

In the study the data were analysed according to multiple criteria:

- 'Self' or 'Observer' position
- Gender
- Nationality
- Level of management experience
- Age.

Certain effects that are commonly observed in leadership questionnaires – for example, observers tend to give lower scores, and women tend to receive higher scores on questions concerning emotions and feelings – were also apparent in the GlobeInvent. The effect of experience was clearly noted in the comparison of the MBA population with the executive population in the database. It was interesting to observe the lack of an age effect concerning the various dimensions. Because the sample sizes were rather low, this effect is worth further exploration to see if there are other ways to explain the lack of difference apart from efficacy in selection.

The validity of the GlobeInvent is further confirmed by the instrument's sensitivity in comparing various leadership dimensions using the multiple criteria noted above. A more in-depth study of the nature of the interactions between 'Self' and 'Observer' scores for certain variables (especially gender and nationality) would be worth undertaking with a larger database, particularly given that this kind of analysis has not been done for 360-degree feedback instruments.

Future considerations

Although, at first step, venturing into the domain of leadership research may seem like walking on quicksand, the make-up of the *Global Leadership Life Inventory* and the findings of this investigation of that instrument demonstrate the considerable advances that have been made over recent decades. Due to promising new research directions, the prevailing attitude of disillusionment with leadership studies is shifting in a more positive direction. Most researchers of leadership now perceive the importance of studying leaders in their 'natural' setting. As a result, there is now considerable agreement about the roles effective leaders play.

The dimensions that are reflected in the GlobeInvent – dimensions that are deemed to be important for exemplary leadership – are grounded in the personal experiences of a large number of very senior executives. These dimensions can now be assessed using the GlobeInvent. The objective of this questionnaire is to deepen test-takers' awareness of the importance of the various dimensions and to encourage test-takers to undertake a close examination of their own capabilities in each of these critical areas. Comparing 'Self' scores with the scores given by 'Observers' can be the beginning of a life-changing journey of self-discovery. Self-exploration matters: not only does it have a profound effect on one's own behaviour, but it affects the lives of others. In the case of senior executives, whose decisions impact on thousands, it affects overall organizational functioning.

As the writer Sholem Aleichem once said, 'If somebody tells you that you have ears like a donkey, pay no attention. But if two people tell you, buy yourself a saddle.' Leaders who fail to take their irrational side into account are like captains who blindly plough their ship into a field of icebergs: the greatest danger is hidden below the surface.

References

Alimo-Metcalfe, B. and Alban-Metcalfe, J.R. (2001) 'The Development of a New Transformational Leadership Questionnaire', *Journal of Occupational and Organizational Psychology*, 74: 1–27.

Arbuckle, J.L. and Wothke, W. (1999) *Amos 4.0 User's Guide*. Chicago, IL: Small Waters Corporation.

Argyris, C. and Schön, D. (1978) *Organizational Learning: A Theory of Action Perspective*. Reading, MA: Addison-Wesley.

Bass, B.M. (1990) *Bass & Stogdill's Handbook of Leadership: Theory, Research, and Applications*. New York: The Free Press.

Bass, B.M. and Avolio, B.J. (1990a) *The Multifactor Leadership Questionnaire*. Palo Alto, CA: Consulting Psychologists Press.

Bass, B.M. and Avolio, B.J. (1990b) *Transformational Leadership Development: Manual for the Multifactor Leadership Questionnaire*. Palo Alto, CA: Consulting Psychologists Press.

Bass, B.M. and Avolio, B.J. (1993) *Transformational Leadership: A Response to Critiques in Leadership Theory and Research Perspectives and Directions*. New York: Academic Press.

Bennis, W. and Nanus, B. (1985) *Leaders: The Strategies for Taking Charge*. New York: Harper & Row.

Browne, H.W. and Cudeck, R. (1993) 'Alternative Ways of Assessing Model Fit'. In Bollen, K.A. and Long, J.S. (eds) *Testing Structural Equations Models*. Newbury Park, CA: Sage Publications, pp. 136–62.

Cohen, J. (1992) 'A Power Primer', *Psychological Bulletin*, 112: 155–9.

Corroyer, D. and Rouanet, H. (1994) 'Sur l'importance des effets et de ses indicateurs dans l'analyse statistique des données', *L'Année Psychologique*, 607–24.

Cortina, J.M. (1993) 'What Is Coefficient Alpha: An Examination of Theory and Application', *Journal of Applied Psychology*, 78: 98–104.

Dalessio, A.T. (1998) 'Using Multi-Source Feedback for Employee Development and Personnel Decisions'. In Smither, J.W. (ed.) *Performance Appraisal: State-of-the-Art in Practice*. San Francisco, CA: Jossey-Bass, pp. 278–330.

Devereux, G. (1978) *Ethnopsychoanalysis*. Berkeley and Los Angeles, CA: University of California Press.

Erikson, E.H. (1958) *Young Man Luther*. New York: Norton.

Erikson, E.H. (1969) *Gandhi's Truth: On the Origins of Militant Nonviolence*. New York: Norton.

Geertz, C. (1973) *The Interpretation of Culture*. New York: Basic Books.

Glaser, B.G. and Strauss, A.L. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Hawthorne, NY: de Gruyter.

Hazucha, J.F., Hezlett, S.A. and Schneider, R.J. (1993) 'The Impact of 360-Deggree Feedback on Management Skills Development', *Human Resource Management*, 32: 325–51.

House, R.J. and Aditya, R.N. (1997) 'The Social Scientific Study of Leadership: Quo Vadis?', *Journal of Management*, 23(3): 409–73.

Jöreskog, K.G. and Sörbom, D. (2000) *LISREL 8.5 User's Reference Guide*. Chicago, IL: Scientific Software International.

Kets de Vries, M.F.R. (1993) *Leaders, Fools, and Impostors*. San Francisco, CA: Jossey-Bass.

Kets de Vries, M.F.R. (2001) *The Leadership Mystique*. London: The Financial Times/Prentice Hall.

Kets de Vries, M.F.R. and Florent-Treacy, E. (1999) *The New Global Leaders: Percy Barnevik, Richard Branson, and David Simon and the Making of the International Corporation*. San Francisco, CA: Jossey-Bass.

Kluger, A.N. and DeNisi, A. (1996) 'The Effects of Feedback Interventions on Performance: A Historical View, a Meta-Analysis, and a Preliminary Feedback Intervention Theory', *Psychological Bulletin*, 119: 254–84.

Kohut, H. (1977) *The Restoration of the Self*. Madison, CN: International Universities Press.

Kohut, H. (1985) *Self Psychology and the Humanities*. New York: Norton.

Kouzes, J.M. and Posner, B.Z. (1995) *The Leadership Challenge*. San Francisco, CA: Jossey-Bass.

London, M., Wohlers, A.G. and Gallagher, P. (1990) '360-Degree Feedback Surveys: A Source of Feedback to Guide Management Development', *Journal of Management Development*, 9: 17–31.

Nunnally, J.C. (1978) *Psychometric Theory*. New York: McGraw-Hill.

Peterson, R.A. (1994) 'Cronbach's Alpha Coefficient: A Meta-Analysis', *Journal of Consumer Research*, 21: 381–91.

Posner, B.Z. and Kouzes, J.M. (1988) 'Development and Validation of the Leadership Practices Inventory', *Educational and Psychological Measurement*, 48: 483–96.

Posner, B.Z. and Kouzes, J.M. (1993) 'Psychometric Properties of the Leadership Practices Inventory-Updated', *Educational and Psychological Measurement*, 53: 191–9.

Schein, E.H. (1987) *The Clinical Perspective in Fieldwork*. Newbury Park, CA: Sage.

Stogdill, R.M. (1948) 'Personal Factors Associated with Leadership: A Survey of the Literature', *Journal of Psychology*, 25: 35–71.

Stogdill, R.M. (1974) *Handbook of Leadership*. New York: The Free Press.

Tichy, N.M. and Devanna, M.A. (1986) *The Transformational Leader*. New York: Wiley.

Van Maanen, J. (1988) *Tales of the Field*. Chicago, IL: University of Chicago Press.

Van Velsor, E. and Leslie, J.B. (1991a) *Feedback to Managers*, Vol. 1. *Guide to Evaluating Multirater Feedback Instruments*. Greensboro, NC: Center for Creative Leadership.

Van Velsor, E. and Leslie, J.B. (1991b) *Feedback to Managers*, Vol. 2. *A Review and Comparison of Sixteen Multirater Feedback Instruments*. Greensboro, NC: Center for Creative Leadership.

Walker, A.G. and Smither, J.W. (1999) 'A Five-Year Study of Upward Feedback: What Managers Do with their Results Matters', *Personnel Psychology*, 52: 393–423.

Yukl, G. (1994) *Leadership in Organizations*. Englewood Cliffs, NJ: Prentice Hall.

Zaleznik, A. (1977) 'Managers and Leaders: Are they Different?', *Harvard Business Review*, May–June: 47–60.

Zaleznik, A. and Kets de Vries, M.F.R. (1975) *Power and the Corporate Mind*. Boston, MA: Houghton Mifflin.