JOURNAL OF APPLIED
COMPUTER SCIENCE
Vol. xx. No xx (200x), pp. xx-xx

ABOUT BUSINESS DECISION MAKING BY A CONSISTENCY-DRIVEN PAIRWISE COMPARISONS METHOD

Waldemar W. Koczkodaj[1], Rolland LeBrasseur[2], Aleksander Wassilew[3], Ryszard Tadeusziewicz[4]

[1]*Laurentian University*
Department of Mathematics and Computer Science
Ramsey Lake Road Sudbury, Ontario
Canada P3E 2C6
wkoczkodaj@cs.laurentian.ca

[2]*Laurentian University*
Dept of Commerce
Ramsey Lake Road Sudbury, Ontario
Canada P3E 2C6
RLebrass@laurentian.ca

[3]*Warsaw School of Economics*
Department of Business Informatics
Al. Niepodleglosci 164
02-554 Warsaw, Poland
alex@sgh.waw.pl

[4]*AGH University of Science and Technology*
Department of Automatics
Al. Mickiewicza 30
30-059 Kraków, Poland
rtad@agh.edu.pl

Abstract. *Writing this paper has been inspired by the most recent economic crises in the world. Better decision making methods are more needed now than ever before. This study presents an innovate approach to the assessment of management capability in businesses. It is based on the consistency-driven pairwise comparisons method. A proposed conceptual model of performance is flexible and adaptable to different requirements and preconditions (e.g., grant or loan applications). Considering the complexity, a hierarchical structure is used and an inconsistency analysis is performed for all the levels of the structure. The pairwise comparisons method synthesizes together performance assessments assessed at two levels (in our case; there may be more levels in general). The method of consistency-driven pairwise comparisons can be combined with other quantitative and qualitative assessment methods (including brainstorming and Delphi method). Non measurable criteria which often bypassed in other approaches, can be included in the presented model. The consistency-driven pairwise comparisons method contributes to the reliability of assessment through the consistency analysis and solid statistical studies show the*

accuracy improvement. The management capability merit index (MC-merit) and a procedure for computing it are introduced.
Keywords: *pairwise comparisons, inconsistency analysis, expert opinion, management capability, business performance, assessment.*

1. Introduction

The most recent economic crises in the world in other industrialized countries clearly indicate that survival and growth of businesses depends of better decision making. Small businesses have attracted national attention because these enterprises make a significant contribution to economic development (shown in [1]) and will be used as a demonstration case here. The presented methodology is general enough for any business size. In particular, our approach can be easily customized to the mortgage evaluation which was at the core of the most recent economic crises. However, it would require a more specific knowledge and as such could be less suitable for the presentation of the method.

A statistical report [2] by the Organization for Economic and Cooperative Development documented that in 1995 the service sector represented 67.7% of civilian employment in the G7 countries and 64.5% in the 15 countries of the European Union. By 2005, these figures had increased to 73.7% and 69.8%, respectively.

For our purposes, we emphasize that both theoretical approaches link firm success to management capability (see [3]). In Canada, governments have numerous financial support programs for startup and expansion. In order to qualify for a program, explicit criteria must be matched.

Table 1. Explicit Criteria Used for Small Business Support Programs

Number of employees
Gross average sales
Independence of owner(s) from outside control
No dominance in the market served
Owner(s) actively involved in management

Table 1 summarizes criteria most frequently cited for qualifications for business support programs. If the applicant/business satisfies these criteria, the government agent (or banking intermediary) is left with the task of assessing a number of intangible factors, among them management capability. Because of the high failure rates among small businesses, the assessment of this intangible becomes mandatory as explained in [4].

Management capability may be defined as competence in a wide range of functional areas such as financial, marketing, human resource and operational management, as well general management practices such as planning and administration. The expectation is that higher management capability leads to better overall business performance. This higher level of performance is expected to be reflected in higher sales, profits, and cash flows. Table 2 summarizes the various factors comprising management capability. This summary reflects the management academic literature and its interest in the industrial and managerial experience of the business founder and/or partners, and their networking in the business community. Management capability creates the human, financial, physical and information resources which make the business successful for the foreseeable future.

Table 2. Four pillars of the management capability assessment

Financial and Operations Capability (FOC)		
F1	Finance	Financial management
F2	Margin	Pricing for competitiveness and margin
F3	Oper1	Operations (inventory control, scheduling)
F4	UtBgPln	Utilizing budgets for planning and control
General Management Capability (GMC)		
F5	BsnPln	Busines

		s planning
F6	Training	Training and supervision of employees
F7	Advert	Advertising and promotion
F8	CusServ	Customer service
F9	NewProd	Developing and introducing new products
F10	Image	Business image and visibility
Technological Capability (TC)		
F11	BsnSys	Use of compute

		r technology for business systems
F12	eComm	Use of computer technology for e-commerce
F13	Oper2	Use of computer technology for operations
Strategic Management Capability (SMC)		
F14	MnChg	Management of change
F15	Export	Export-orientation and networ

		king
F16	MnTeam	Management team

Table 2 organizes the various management capabilities into the Four Pillars. The first pillar is Financial and Operations Capability and reflects the operational core of the enterprise (as explained in [5]). The second pillar, General Management Capability, covers the human factors including employees and customers, as well as general planning and new products (see [6]). The third pillar is Technical Capability reflecting the use of computer technology for administrative systems, operations, and e-commerce (as addressed in [7]). The fourth and last pillar, Strategic Management Capability, covers the breadth of top management skills to enact change and engage with the competitive environment (see [8] for details).

2. Preliminaries of the management capability assessment

From the perspective of awarding grants or loans, evaluating the management capability is important yet no comprehensive solution has been proposed or at least such solution has never been popularized. Most assessment methodologies are based on *return-on-investment* (ROI), economic activities, deterministic econometric methods, models, multivariate models, and preference mapping. All the above are precise and powerful tools, but all require precise data which are rarely available and, as we can see, take into account only the economic perspective.

The above remarks are crucial for understanding the novelty of the consistency-driven pairwise comparisons approach to the management capability assessment. Nothing can substitute for precise methods (such as the above-mentioned methodologies), but in the absence of precise data a second-best alternative, *expert assessment*, is not only a vital alternative but it may turn the only solution. Our approach helps to assess business performance in situations where precise data is unavailable or too time consuming to collect. The proposed method is particularly useful for assessing when a database of small firms is maintained. It often takes place in case of bidding. When both precise and imprecise data that are available, they can be combined with the expert judgment of assessing specialists who have been assigned to assess the firm by, for example adding one more level to our model: with the groups labeled by: QD (quantifiable data) and NQD (non quantifiable data).

The method of consistency-driven pairwise comparisons analyzes inconsistencies of the subjective assessments provided by experts by comparing two criteria at a time as explained in Appendix B. A case presented in this paper includes (but is not limited to) the financial, marketing, human resource and operational management. In the management capability evaluating process, factors with about the same degree of importance must be considered simultaneously, and assigning them precise weights is also important. Weights of evidence are computed on the basis of partial judgments expressed by comparing assessment criteria in pairs. Such assessments are analyzed for consistency and are enhanced with the help of software (the Concluder program, released to public domain, runs under Windows XP on a personal computer).

MC-merit index (*MC-merit*) can be used for assessing individual business firms. *MC-merit* combines different aspects of management capability. Using a structured approach to the conceptual model of assessment is expressed by the hierarchy in the model (multi level approach) and the use of consistency-driven pairwise comparisons for evaluating criteria enhances precision. The method of consistency-driven pairwise comparisons is applied to the assessment criteria in pairs, as it is easier and more precise than looking at all criteria at once. In particular, the localization of the most inconsistent assessments as introduced in [9], helps to refine assessments expressed by experts.

One needs to reflect on the measurement process in general to understand the essence and potential gain in the precision by pairwise comparisons relative to direct estimation (for example, by a measurement or judgment). There is no standard measure (such as the cubic meter) for assessing the business performance. The lack of standard measure forces us to compare one object to another. Consider this example where two stones, A and B, can be weighed if we have a scale. When a scale is not accessible, we often *weigh* them each in one hand. Without using a standardized measuring device, it is easier to say that A is 1.5 times heavier than B rather than to guess the exact weight of each stone. Interestingly enough, the use of a standardized measure such as a kilogram is also a pairwise comparison. The statement the weight of A is 2.5 kilograms" is an abbreviation of" *by a pairwise comparison of A to one kilogram we have a factor of 2.5*.

A Monte Carlo experiment with bars of randomly generated lengths showed a considerable relative improvement of accuracy in estimation of their lengths by using pairwise comparisons from about 15% error to 5% error was observed and was verified statistically in [10]. In this Monte Carlo experiment, respondents were first asked to estimate the lengths directly. Subsequently, respondents were asked to do the same by comparing bars in pairs. Not much change in approach but the relative error improvement was 300% (that what we get from 15% divided by 5%) and can be translated into millions of dollars saved for mistakes (awarding a grant or loan to a company later on proclaiming bankruptcy or declining a loan to a company capable to survive with the loan/grant).

We have become so accustomed to having standards that sometimes it is difficult to imagine a situation where no standard measure exists. Management capability evaluating are commonly based on non-measurable criteria such as design, reliability and experience. Comparing assessment criteria in pairs is the key issue and the solution to our problem. Condorcet was probably the first researcher who used the pairwise comparisons in 1785 for a voting method in [11]. However, it was Fechner who in 1860 described the pairwise comparisons method in [12] but from the psychometric perspective only. Thurstone in [13] not only described pairwise comparisons method but for the first time has provided a statistical analysis and proposed a solution. In [14], Saaty introduced a hierarchy, which was instrumental for practical applications, and eigenvalue-based inconsistency which was only a global indicator and as such could not localize the most inconsistent elements as it can be done by the inconsistency definition proposed in [9]. Both inconsistencies were recently analyzed in [15]. Monte Carlo study [10] provided evidence that for small inconsistencies both the geometric means solution (used in this study) and the eigenvector solution (as Saaty proposed in [14] as identical from statistical point of view.

The consistency-driven approach incorporates the reasonable assumption that by finding the most inconsistent assessments, one is able to reconsider his/her own opinions. The identification of inconsistencies is done by software (the Concluder program) which highlights criteria requiring reconsideration. This in turn contributes to improvements in the accuracy of assessments. A dynamic process of consistency analysis is facilitated by the software, which displays the most inconsistent assessments in a contrasting color on the computer screen as demonstrated in the next section. The procedure usually starts (after an appropriate feasibility study and data gathering, which are not addressed here) with a listing of all possible criteria. In our case, the criteria mentioned in the preceding section are used.

3. A conceptual model of the assessment process

The next step in the method is to group criteria together. A rule of thumb proposed by Saaty in [14] is that no group should have more than seven criteria. A larger number of criteria in one group is impractical because the number of all combinations of pairs grows rapidly (for seven criteria, it is 21). A practical solution for groups with more criteria is to split them into subgroups which create the next level. If there are, for example, 30 criteria, splitting them into five groups of six is as good a solution as a split into: 4, 5, 3, 7, 6, and 5, or any other arrangement. It is important to group criteria on the basis of natural affinity. A purely mechanical split (for example, by alphabetical order) is not recommended, as it may be much harder to compare such criteria on a pair-by-pair basis. The presented model has been simplified for the presentation purpose. The method, however, is general, and is expandable to other types of business performance without any additional theoretical pitfalls. The criteria are divided into financial, operational, marketing and human resource management criteria which are used to create the hierarchical model shown in Fig. 1.

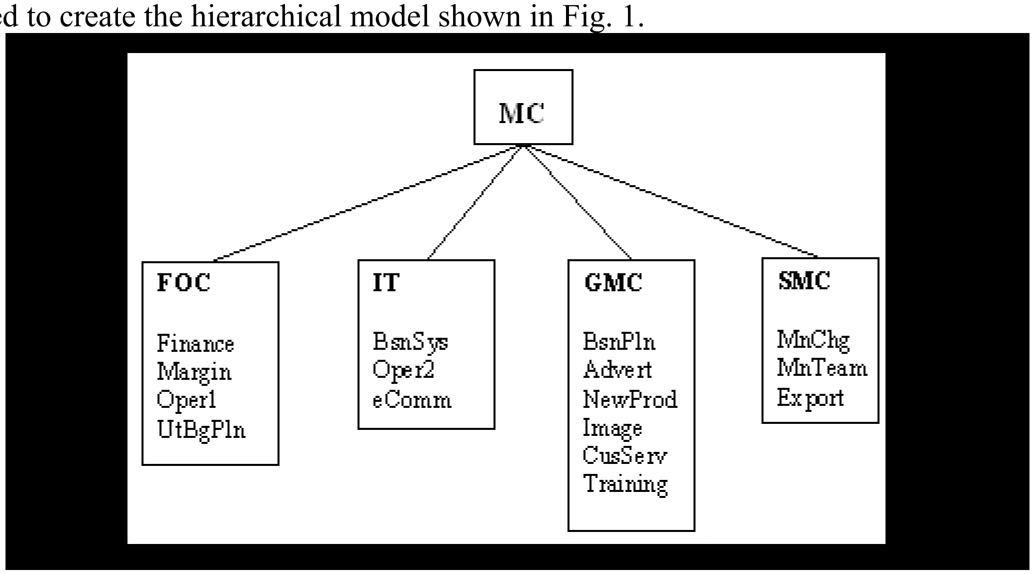

Fig.. 1. The Conceptual model of the assessment process

4. Development of weights and interpretation of obtained results

After a conceptual model of assessment is created, the consistency-driven pairwise comparisons method is used to compare groups of criteria and criteria within individual groups by using a rather rough scale from 1 to 5 presented in Table 2. A scale of 1 to 5 (Table 2) is used to compare criteria in pairs for expressing an expert's preferences. The reasoning for the selection of the scale is outside the scope of this paper, but it should be noted that for precise measurements one can use intermediate values. for example, 3.142; the use of a wider scale for subjective assessments may only confuse the user. Based on our experience with the presented method, a scale of 1 to 5 is effective for the consistency-driven pairwise comparisons.

The challenge posed to the pairwise comparisons method comes from the lack of consistency in assessments which arise in the real world. The basics of consistency analysis are shown in Appendix B (the successful use of the method does not require an understanding of all mathematical aspects of the consistency analysis, but certainly this can contribute to the increased confidence in the final results). Although the consistency analysis may seem complicated, the software developed for this analysis makes it easy to use (most of our users consider it equivalent to playing a video game). During the analysis, the most inconsistent combinations of criteria are displayed visually. By decreasing or increasing a previously entered value (by a relative comparison of two criteria), the user develops an intuition for what should be changed to achieve an acceptable level of consistency. Changing the numbers simply to achieve lower inconsistency is not advisable. Leaving the highly inconsistent assessments unchanged may in fact be a better alternative than trying to please the system by entering meaningless numbers. High inconsistency indicates that the user should reconsider the comparisons of the three criteria that contribute most to the inconsistency. In some cases, changes will need to be extensive and the entire model may need revamping. This may require refinements of the definitions of the criteria, adding new criteria, or dropping some of the criteria. It is an interactive process requiring input from knowledgeable assessment specialists. In practice, inconsistent assessments are unavoidable when at least three factors are independently compared against one another (see Appendix B for more explanations; in cases involving only two criteria inconsistency does not occur, but only inaccuracy results). All of the above computations, including the final weights, are done automatically by the software. It is not important to address all mathematical aspects of obtaining the final weights (see Appendices A and B).

The main goal of this approach is to obtain assessments of a business performance. One of the initial steps is to assign numerical values to the entire list of criteria (Fig. 1) using past experience, intuition, and common sense. With numerical values assigned to each criterion, one can easily construct the pairwise comparisons matrix A (see Appendix A) by simple division of the assigned numerical values with the corresponding criteria. In our case, we may assign numerical values to all twenty one criteria, for example, 20 to t*raining of employees* and 8 to *general management*. The assigned numerical values can be used to establish an initial value of the relative comparison of *training of employees* against *general management* as 2.5 (that is 20/8). After a careful examination of the matrix created with these quotients we may, for example, discover that certain ratios are too small, whereas others are beyond common sense. One may create inconsistencies by changing the ratios, which can be handled by the method, as explained later in this section. The computer program locates and highlights the most inconsistent assessments, which can be improved by re-examining the criteria.

Table 3. Comparison scale

Code	Definition of intensity or importance	Explanation
1	Equal importance	Two criteria equally contribute to the objective
2	Weak importance	Experience and judgments slightly favor one criterion over another
3	Essential or strong importance	Experience and judgments favor one criterion over another
4	Demonstrated importance	The criterion is strongly favored and its dominance is demonstrated in practice
5	Absolute importance	The highest affirmation degree of favoring one criterion over another
2.4, etc.	Intermediate judgments	When compromise is needed

We can begin by examining criteria in the FOC group (Fig. 2). There has been shown two screens with relative comparisons forming initial and enhanced matrices. Scores in the most inconsistent triad are highlighted by the software. In the example, we have assigned the relative importance of 0.5, which is less than *equal* importance according to Table 3, to *Margin* when compared to UtBgPln (utilizing budgets for planning and control). This assigned value is for illustrative purposes only, as the software can process any reasonable preferences. *Finance* in comparison to *Oper1* (operations, inventory control, scheduling) has been assigned a value of 2 (*moderate importance*). The initial values for *Finance* against *UtBgPln* is only 1 (*equal importance*), but for *Margin* against *Oper1* were given a value of 3 (*strong importance* according to Table 3). The initial evaluating of importance of *Oper1* against *UtBgPln* was set to *moderate importance* (of one criterion over another and is reflected by 2 in the last column of row three (left image in Fig. 2). The maximal inconsistency computed for the above comparisons is 0.50 and it is exceeding 0.33 (a threshold as explained in [9] indicates that relative comparisons of the criteria in this group are internally inconsistent, which reflects a tendency of human nature to overestimate criteria along the line: *everything is more important than anything else*. After applying the consistency analysis (Appendix B), the relative comparisons were refined after the software has identified the most inconsistent triad of the comparisons. In the presented case, it was easy to reduce inconsistency through modifying just 1 comparison – for *Margin* against *UtBgPln* from 0.5 to 0.75, what has proved goodness for initial assessments. The new inconsistency index was computed as 0.33 assumed to be an acceptable threshold of inconsistency as introduced in [9] and addressed in Appendix B.

Fig. 2. Initial (left) and enhanced (right) relative evaluating for *FOC* group of criteria

	Finance	*Margin*	*Oper1*	*UtBgPln*
Finance	**1.00**	**1.00**	**2.00**	**1.00**
Margin		**1.00**	**3.00**	**0.50**
Oper1			**1.00**	**2.00**
UtBgPln				**1.00**

Maximum inconsistency: 0.50

	Finance	*Margin*	*Oper1*	*UtBgPln*
Finance	**1.00**	**1.00**	**2.00**	**1.00**
Margin		**1.00**	**3.00**	**0.75**
Oper1			**1.00**	**2.00**
UtBgPln				**1.00**

Maximum inconsistency: 0.33

Values below the main diagonal do not need to be entered by the user in the software system. They are reciprocal to the corresponding values in the upper triangle, that is x^{-1} for x in the upper triangle.

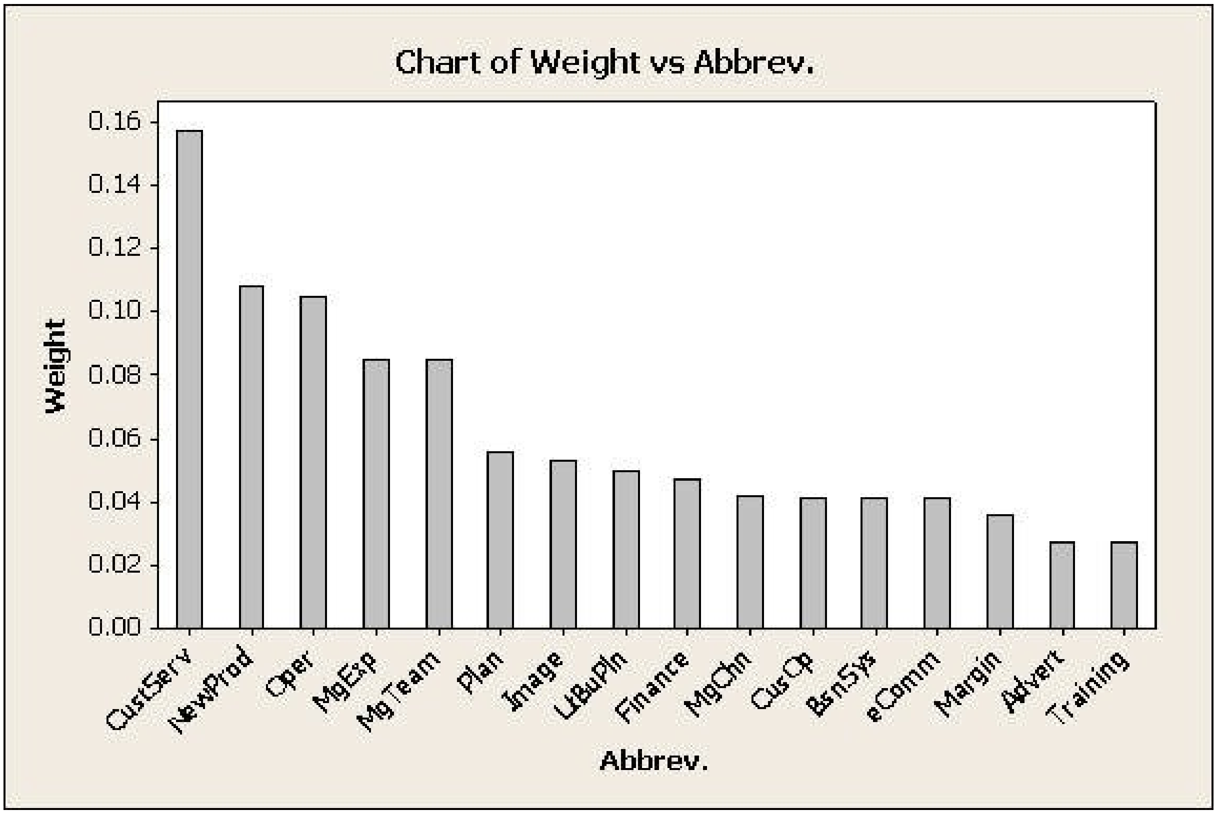

Fig. 3. Final weights of assessed criteria

When an acceptable consistency level is achieved, the weights derived by the software for all criteria (shown in Fig. 3) are used to compute the *MC-merit* indexes for a grant or loan applications (presented in Table 4 and Figure 4). Evaluation scores between 0 and 5, with 0 expressing the lowest and 5 the highest value, are assigned to each criterion by the specialist on the basis of knowledge of the case. The evaluation scores are then multiplied by the weights, which are supplied by the consistency-driven analysis, for each criterion; the *MC-merit* indexes which are then computed for each case are displayed on the computer screen. These are multiplied by the respective weights and divided by 5 (which is the maximum score) to produce a total assessment score. Consistency analysis (Appendix B) and refinement of assessments are supported by the user-friendly software system.

Table 4. Evaluations of three example companies, C1, C2, and C3

Cr	Abbrev.	Criterion name	Weight	C1	C2	C3
C1	CustServ	Customer service	15.7%	4	4.5	5
C2	NewProd	Developing and introducing new products	10.8%	5	3	2
C3	Oper	Operations (inventory control, scheduling)	10.5%	3	5	4.5
C4	MgExp	Export-orientation and networking	8.5%	2	5	2
C5	MgTeam	Management team	8.5%	2	4	3
C6	Plan	Business planning	5.6%	4	3.5	5
C7	Image	Business image and visibility	5.3%	4	2.5	3
C8	UtBuPln	Utilizing budgets for planning and control	5.0%	2	3	5
C9	Finance	Financial management	4.7%	3.5	4	2
C10	MgChn	Management of change	4.2%	4	2.5	1
C11	CusOp	Use of computer technology for operations	4.1%	2.5	3	1.5
C12	BsnSys	Use of computer technology for business systems	4.1%	4	5	3
C13	eComm	Use of computer technology for e-commerce	4.1%	5	2	2
C14	Margin	Pricing for competitiveness and margin	3.6%	4	4	4
C15	Advert	Advertising and promotion	2.7%	3	5	2
C16	Training	Training and supervision of employees	2.7%	5	3	5
		Totals	100.0%	3.52	3.86	3.32

5. Conclusions

A consistency-driven pairwise comparisons method is based on a logical model for assessing management capability. The model can utilize the in-house expertise and published information on the past projects. It also can be supported by any other techniques, such as a brainstorming, or by the *Delphi method* which is a systematic interactive forecasting method for obtaining forecasts from a panel of independent experts.

The business performance may be modeled in many different ways. However, the identification of major criteria is an essential component in building each model. Once this is done, the final weights can be easily computed from the relative pairwise comparisons. The model demonstrated in this paper has a generic character and has been constructed for reason of illustration. Nevertheless, the presented set of criteria can be considered at least as reasonable. If needs come, they can be easily altered to a useful form by inclusion of other criteria or exclusion of irrelevant criteria. The outlined methodology is flexible and allows consideration of all criteria at hand, including both quantitative and qualitative factors.

This is just the beginning in establishing interest and necessary direction towards solving problems of business management capability assessment. While the pairwise comparisons method has been used for 223 years (Condorcet made a reference to it in 1785, the time of the French Revolution, [11])), to our knowledge, it has not been used to improve the business management capability assessment process at a satisfactory level. Although it is a time consuming exercise, it can be of global benefit to the rest of medical community and as such it may be worth trying. A jury of experts from the academia, industry, and government may be called to set the priorities and a method presented in [16] can be used in case of ties of expert opinions.

Acknowledgment

This work has been supported by NSERC, Ministry of Northern Development and Mines (for developing software for Abandoned Mines Hazard rating), and Laurentian University Research Grant in Canada.

References

[1] OECD (2000), *High Growth SMEs: Phase II Synthesis Report*.

[2] OECD (2007). Organization for economic and cooperative development. *OECD in Figures 2006-2007.* OECD Observer, 2006/Supplement 1, p. 32.

[3] Boeker W., Wiltbank R. (2005) New Venture Evolution and Managerial Capabilities. Organization Science, 16(2), 123-133.

[4] Jungman H., Okkonen J, Rasila T, Seppai M. (2004) Use of performance measurement in V2C activity. Benchmarking, 11(2), 175-189.

[5] Wood EH (2006). The internal predictors of business performance in small firms: A logistic regression analysis. Journal of Small Business and Enterprise Development, 13(3), 441-453.

[6] Ibrahim N.A., Angelidis J.P., Faramarz P. (2004) The Status of Planning in Small Businesses. American Business Review, 22(2), 52.

[7] Chong S., Pervan G. (2007) Factors Influencing the Extent of Deployment of Electronic Commerce for Small-and Medium-Sized Enterprises. Journal of Electronic Commerce in Organizations, 5(1), 1-29.

[8] Ardichvili A., Cardozo R., Reynolds P., Williams M. (1998) The New Venture Growth: Functional Differentiation and the Need for Human Resource Development Interventions. Human Resource Development Quarterly, 9(1), 55-70.

[9] Koczkodaj W.W. (1993) *A New Definition of Consistency of Pairwise Comparison*. Mathematical and Computer Modelling, 18(7), 79-84.

[10] Herman M., Koczkodaj W.W. (1996) Monte Carlo Study of Pairwise Comparisons), Information Processing Letters, 57(1), pp. 25-29.

[11] Condorcet M. (1785) The Essay on the Application of Analysis to the Probability of Majority Decisions, Paris: Imprimerie Royale.

[12] Fechner G.T. (1860) Elements of PsychoPhysics, Vol. 1, New York: Holt, Inehart & Winston, 1965, translation by H.E. Adler of Elemente der Psychophysik), Leipzig: Breitkopf und Hartel.

[13] Thurstone L.L. (1927) *A Law of Comparative Judgments*, Psychological Reviews, Vol. 34, 273-286.

[14] Saaty T.L. (1997) *A Scaling Method for Priorities in Hierarchical Structure*. Journal of Mathematical Psychology, 15(3), 234-281.

[15] Bozóki, S., Rapcsák T., (2007) On Saaty's and Koczkodaj's inconsistencies of pairwise comparison, Journal of Global Optimization, 2007 (available online at

http://www.springerlink.com/content/v2x539n054112451/, the hard copy due soon)
[16] Janicki R., Koczkodaj W.W. (1996) *A Weak Order Approach to Group Ranking*, Computers and Mathematics with Applications, 32(2), pp. 51-59.
[17] Chrisman J.J, Bauerschmidt A., Hofer C.W. (1999) "The Determinant of New Venture Performance: An Extended Model," *Entrepreneurship Theory & Practice* 23(1), 5-29.
[18] Industry Canada, (2008). *Key Small Business Statistics*, www.strategis.gc.ca/sbstatistics, information retrieved March 16, 2008
[19] Koczkodaj W.W. (1996) *Statistically Accurate Evidence of Improved Error Rate by Pairwise Comparisons*. Perceptual and Motor Skills, 82, 43-48.

Appendix A - Basic Concepts of Pairwise Comparisons

A n by n pairwise comparisons matrix is defined as a square matrix $A=[a_{ij}]$ such that $a_{ij}>0$ for every $i,j=1,...,n$. Each a_{ij} expresses a relative preference of criterion (or stimulus) s_i over criterion s_j for $i,j=1,...,n$ represented by numerical weights (positive real numbers) and w_i and w_j respectively. The quotients $a_{ij}= w_i / w_j$ form a pairwise comparisons matrix A.

Matrix A:

A=

1	a_{13}	...	a_{1n}
$1/a_{13}$	***1***	...	a_{2n}
...	...	...	...
$1/a_{1n}$	$1/a_{2n}$	...	***1***

A pairwise comparisons matrix $\boldsymbol{A}$ is called *reciprocal* if $a_{ij}=1/a_{ji}$ for every $i,j=1,...,n$ (then automatically $a_{ii}=1$ for every $i=1,...,n$ because they represent the relative ratio of a criterion against itself). A pairwise comparisons matrix $\boldsymbol{A}$ is called *consistent* if $a_{ij} \cdot a_{jk}=a_{ik}$ holds for every $i,j,k=1,...,n$ since $w_i /w_j \cdot w_j /w_k$ is expected to be equal to w_i /w_k. Although every consistent matrix is reciprocal, the converse is not generally true. In practice, comparing of s_i to s_j, s_j to s_k, and s_i to s_k often results in inconsistency amongst the assessments in addition to their inaccuracy; however, the inconsistency may be computed and used to improve the accuracy.

The first step in pairwise comparisons is to establish the relative preference of each combination of two criteria. A scale from 1 to 5 can be used to compare all criteria in pairs. Values from the interval $[1/5,1]$ reflect inverse relationships between criteria since $s_i/s_j=1/(s_j/s_i)$. The consistency

driven approach is based on the reasonable assumption that by finding the most inconsistent judgments, one can then reconsider one's own assessments. This in turn contributes to the improvement of judgmental accuracy. Consistency analysis is a dynamic process which is assisted by the software.

The central point of the inference theory of the pairwise comparisons is Saaty's Theorem presented in [14]. It states that for every n by n consistent matrix $A=[a_{ij}]$ there exist positive real numbers $w_1,...,w_n$ (weights corresponding to criteria s_1,.$_n$.. ,s) such that $a_{ij}=w_i/w_j$ for every $i,j=1,...,n$. The weights w_i are unique up to a multiplicative constant. One possible solution to the weights problem is the least

squares

solution, but it was far more computationally demanding than a method of row geometric means. A statistical experiment demonstrated that the accuracy, that is, the distance from the original matrix A and the matrix A ▪ reconstructed from weights with elements $[a_{ij}$

$]=[w_i/w_j]$, does not strongly depend on the method. There is, however, a strong relationship between the accuracy and consistency.

An important problem is how to begin the analysis. Assigning weights to all criteria (e.g., *A=18*, *B=27*, *C=20*, *D=35*) seems more natural than the above process. In fact it is a recommended practice to start with some initial values. The above values yield the ratios: *A/B=0.67*, *A/C=0.9*, *A/D=0.51*, *B/C=1.35*, *B/D=0.77*, *C/D=0.57.* Upon analysis, these may look somewhat suspicious because all of them round to 1, which is of equal or unknown importance. This effect frequently arises in practice, and experts are tempted to change the ratios by increasing some of them and decreasing others (depending on knowledge of the case). The changes usually cause an increase of inconsistency which, in turn, can be handled by the analysis because it contributes to establishing more accurate and realistic weights. The pairwise comparisons method requires evaluation of all combinations of pairs of criteria, and can be more time

consuming because the number of comparisons depends on n^2 (the square of the number of criteria). The complexity problem has been addressed and partly solved by the introduction of hierarchical structures (see [14]). Dividing criteria into smaller groups is a practical solution in cases in which the number of criteria is large.

Appendix B - Consistency Analysis

Consistency analysis is critical to the approach presented here because the solution accuracy of *not*

so

inconsistent matrices strongly depends on the inconsistency. The consistency

driven approach is, in brief, the next step in the development of pairwise comparisons.

The challenge to the pairwise comparisons method comes from a lack of consistency in the pairwise comparisons matrices which arises in practice. Given an *n* by *n* matrix *A* that is not consistent, the theory attempts to provide a consistent *n* by *n* matrix $\underset{\cdot}{A}$that differs from matrix *A* "as little as possible". In particular, the geometric means method produces results similar to the eigenvector method (to high accuracy) for the ten million cases tested. There is, however, a strong relationship between accuracy and consistency

Unlike the old eigenvalue

based inconsistency, introduced in [14], the triad

based inconsistency locates the most inconsistent triads [14].This allows the user to reconsider the assessments included in the most inconsistent triad. Readers might be curious, if not suspicious, about how one could arrive at values such as *1.30* or *1.50* as relative ratio judgments. In fact the values were initially different, but have been refined and the final weights have been calculated by the consistency analysis. It is fair to say that making comparative judgments of rather intangible criteria (e.g., overall alteration and/or mineralization) results not only in imprecise knowledge, but also in inconsistency in our own judgments. The improvement of knowledge by controlling inconsistencies in the judgments of experts, that is, the *consistency*

driven approach, is not only desirable but is essential.

In practice, inconsistent judgments are unavoidable when at least three factors are independently compared against each other. For example, let us look closely at the ratios of the four criteria *A*, *B*, *C*, and *D* in this square Table B:

Table B for the inconsistency analysis

		A	B	C	D
B=	A	1	2	5	4
	B		1	3→2.5	1.9→2
	C			1	0.7→0.8
	D				1

Suppose we estimate ratios *A/B* as *2*, *B/C* as *3*, and *A/C* as *5*. Evidently something does not add up as *(A/B)* ▪ *(B/C)=2* ▪ *3=6* is not equal to *5* (that is *A/C*). With an inconsistency index of *0.17*, the above triad (with highlighted values of *2, 5*, and 3) is the most inconsistent in the entire matrix (reciprocal values below the main diagonal are not shown). A rash judgment could let us believe that *A/C* should indeed be *6*, but we do not have any reason to reject the estimation of *B/C* as *2.5* or *A/B* as *5/3*. After correcting *B/C* from *3* to *2.5*, which is an arbitrary decision usually based on additional knowledge gathering, the next most inconsistent triad is *(5,4,0.7)* with an inconsistency index of *0.13*. An adjustment of *0.7* to *0.8* makes this triad fully consistent (*5· 0.8* is *4*), but another triad *(2.5,1.9,0.8)* has an inconsistency of *0.05*. By changing *1.9* to *2* the entire table becomes fully consistent. The corrections for real data are done on the basis of professional experience and case knowledge by examining all three criteria involved.

An acceptable threshold of inconsistency is 1/3 which takes place for “one judgment is not more than two grades off the scale 1 to 5” from the value which would make a triad fully consistent. Decreasing the inconsistency to zero may be harmful since we deal with subjective assessments and full consistency may be hard to achieve in practice. Only the high inconsistency is harmful. A very small inconsistency, or zero, may even indicate that the “doctored” data were entered hastily without business analysis of inconsistent assessments.