“Merit measures and validation in employee evaluation and selection”

AUTHORS
Wilson P. R. Malebye
Solly M Seeletse
Mario A. Rivera

ARTICLE INFO

JOURNAL
“Problems and Perspectives in Management”

FOUNDER
LLC “Consulting Publishing Company “Business Perspectives”

© The author(s) 2018. This publication is an open access article.
SECTION 3. General issues in management

Wilson P.R. Malebye (South Africa), Solly M. Seeletse (South Africa), Mario A. Rivera (USA)

Merit measures and validation in employee evaluation and selection

Abstract

Applicants for employment are usually compared subjectively in the selection process, and the selections made are typically not reliable, if only because they are seldom verifiable empirically. The present study describes a process of much more objective selection sequence, one that involves a quantitative/mathematical measure that may be used in selecting a candidate applying for a job, in a process then adds two other independent measures to validate the decision taken. The approach followed is a stepwise combination of StOr methods (Statistics and Operations Research, incorporating SAW, TOPSIS, and WP). In this analysis, SAW (simple additive weighting) is used in the first-cut selection process, and TOPSIS (technique for order preference by similarity to ideal solution) and WP (weighted product) are used to validate selections. A practical exercise was developed from an actual selection problem, part of a real-world recruitment task undertaken in an organization for which the authors consulted, and in which the human resources (HR) department wanted to check if their selection was justifiable, and demonstrably valid. The resulting analytical approach was clearly valid, consistent, reliable, and replicable, and convincing to that HR department, since it considered the determinations made by our system quite satisfactory, while theirs could not stand up to empirical testing or corroboration.

Keywords: candidate selection, HRM, StOr, SW, TOPSIS, WP, MCDM, fuzzy-goal programming, person-job requirement fit, person-job environment fit, decisional complexity.

JEL Classification: C61.

Introduction

The selection of applicants for available job opportunities is most often an ambiguous, even ‘fuzzy’ process incapable of validation (Amid, Ghodsypour and O’Brien, 2007). Customarily, the hiring decision involves an interview panel selecting one applicant for a position for which there are many applicants, with interviews sometimes augmented by personality or competency testing, an ‘in-basket’ job simulation task, and similar standard devices. A common practice is to consider the applicant with the highest qualifications (in their application, résumé, interview, and testing performances) to be the best candidate. Generally, merit refers to the highest qualifications set or the highest marks in the entire complex of on-paper traits, interview outcomes, and testing results. This approach assumes that the attributes used in deriving the germane selection criteria were chosen in ways that were commensurate with their relatively-assigned priority weights.

Another dimension to merit-based selection is the rigor of the candidate’s pertinent experience. A candidate with experience and training denoting the ability to work under severe conditions (such as high stress, or in the face of highly-complex and diverse tasks) tends to be more effective in difficult work environments, better able to handle difficult tasks. In selecting candidates, it is sensible to assume that one who can perform difficult tasks can be trusted to perform these best in similarly difficult circumstances. However, these often intangible traits are difficult to measure accurately and consistently, which adds to the complexities associated with recruitment and selection efforts.

In light of these vexing challenges, common as they are in Human Resource Management (HRM), the authors developed a multi-faceted and multi-attribute decision procedure that entails assigning and assessing various role and task attributes serving as criteria for selection. The challenge then involves determining how these criteria can be applied and evaluated rigorously, so as to select the best-suited candidate from a competitive pool.

Different job attributes are not equally weighted, but rather variable in consequence of assigned rank. In any given hiring situation, some attributes are regarded as more important than others, in a hierarchy of candidate traits with differential weighting consistent with the multiple ranking criteria involved. Candidate selection can therefore be approached as a multiple-criteria decision making problem that is variously affected by quantitative and qualitative factors and entails trade-offs among conflicting criteria. Multi-Attribute, Multi-Objective, or, relatedly, Multiple-Criteria Decision Making (MCDM) techniques can be used to help managers systematically evaluate a set of alternative candidates. As Wang, Hang and Dismukes (2004) remind us, in real-life contexts weighting and ranking differences depend on the perceived importance of the criteria involved, a concern to which we will return in the concluding sections of this study.
Consequently, we have developed a Multiple-Criteria Decision Making (MCDM) system to aid in employee recruitment and selection. This system requires the careful articulation of quantitative and qualitative decision criteria, with explication of the choice of relevant attributes, the prioritization and weighting of these attributes, and the specification of evaluative rules (for instance, selectors for tied or very close final rankings). Candidate selection involving an MCDM system is typically affected by manifold conflicting factors corresponding to desired candidate attributes, attributes whose relative ranking depends on the prior determination of variables to be assessed. These factors invariably lead to combinatorial complexity, in both the specification of selection criteria and in the evaluation entailed by the final ranking of candidates, even under static conditions, with complexity increasing exponentially when conditions or preference mechanisms change (Jadidi, Hong, Firouzi, Yusuff and Zulkifli, 2008).

Since there is almost always a combination of tacit (implicit) and specifiable (explicit) qualifications connected to any actual job, some of which likely pull against one another in actual practice, there is bound to be irreducible vagueness and ambiguity involved. The difficulty of selection problems cannot readily be resolved by deterministic models, since these problems lack sharp demarcation. In these cases, fuzzy logic and fuzzy set theory can be an effective tool to handle uncertainty and help solve the selection problems associated with lack of specifiability (Erol, William and Ferrel, 2003; Holt, 1998; Morlacchi, 1997). *Fuzzy goal programming* has been used successfully in various studies for selection problems with multiple sourcing that involved many objectives (Choo and Wedley, 2004; Golec and Kahya, 2007; Kumar, Vart and Shankar, 2004; Laing and Wang, 1992; Lazarevic, 2001; Lovrich, 2000; Royes, Bastos and Royes, 2003; Sakawa, 2002; Wang, Liou and Hung, 2006; Yaakob and Kawata, 1999). Even before utilization of fuzzy-goal programming in HRM contexts, there were efforts at developing more adaptive screening tools tailored to *person-job environment* fit rather than the rigid expectations of mere *person-job requirement* fit, for instance when recruiting professionals with experience in the disability services field (Wong et al., 1992).

Related to the use of MCDM and MODM in selection processes, is multiple-attribute decision making. MADM, SAW, TOPSIS and WP are effective methods which may be used in tandem to handle complex screening tasks in ways that are both rigorous and capable of validation (Hwang and Yoon, 1987). Ideally, the methods to be implemented should assign a rational (empirically defensible) set of weights to desired position-attributes, in calculating merit-based scores, pursuant to a normative process of selection of the best candidate for the job.

1. **Aim of the study**

The authors formulated the research problem as the empirically-defensible articulation of multiterior attributes for prospective hires, along with corresponding evaluative criteria, so as to select a best-fit candidate using one empirical method and then validate the process by using two other empirical measures, independently applied. Ranking candidates on the basis of multiply-weighted selection criteria runs counter to the practice of using mean scores on the premise of equally-important job-related attributes. Instead, the relative importance of attributes is acknowledged in our system by its reliance on differential weights for various attributes, weights that depend on clearly-delineated though also carefully adapted evaluative criteria.

2. **Decision making**

Making rational decisions in selecting an ‘ideal’ candidate among applicants in a competitive and HR recruitment environment is important. Since ideal solutions rarely obtain, it is equally important to approximate a well-delineated ideal as much as possible. As already suggested, multi-attribute decisional analysis best comports to the complex criteria involved in employment searches, in generating the best alternative among a set of feasible alternatives (Cascio, Outtz, Zedeck and Goldstein, 1991; Chen, 2000; Wang and Lee, 2007; Wei and Chen, 2009). This class of techniques requires that decision makers provide qualitative and/or quantitative assessment baselines for determining the potential contribution to job performance of each factor-linked selection criterion, as well as the entirety of selection criteria in combination, all tied closely to the overall objective(s) of the particular hire (Rosanas and Cugueró, 2012). Consequently, MCDM refers herein to screening, prioritizing, ranking, and selecting from a set of decisional alternatives (also referred to as “candidates” or “actions”) under variable, independent, incommensurate, or conflicting criteria (Fenton and Wang, 2006; Jenkins and Van Kerm, 2006).

Such indeterminate decisional tasks will often result in uncertain, imprecise, indefinite, and subjective decisions, in a decision-making process that can easily become intractable. In other words, decision-making often occurs in a fuzzy-logic context where
the information available to the solution of an insufficiently specified problem is wholly imprecise or uncertain. In this regard, the application of fuzzy set theory to multi-criteria evaluation methods (MCEMs, a rubric inclusive of MCDM and MDOM), under the framework of utility theory, has been shown effective (Chen, Lin and Huang, 2006; Kuo, Tzeng and Huang, 2007). Also indicating the use of both fuzzy-set and MCEM methods, and generally of stochastic methods, in the development of feasible solutions, is the randomness of engendered possibilities, since the outcome of a specific hire based on a particular complex of associated traits or attributes cannot be anticipated with certainty.

The overall utility of decisional alternatives with respect to all selection criteria is often represented by a fuzzy number, cast as a fuzzy utility, often (as in our instance) in connection with fuzzy MCEMs (Hazelrigg, 1996; Iyer and Krishnamurti, 1998; Keeney and Raiffa, 1976; Steuer, 1989). Wang and Lee (2007) indicate that the ranking of alternatives is then to be based on the comparison of their corresponding fuzzy utilities (Cf. Golec and Kahya, 2007).

The general concepts of domination structures and non-dominated solutions play an important role in describing decision problems and decision makers’ revealed preferences in MCDM, for instance in resource-constrained allocation models (Yu, 1985; Ellis and Kim, 2005). Olson (1986) confirms that various multi-tiered analytical approaches have been developed in consequence, as decision aids. However, their efficacy as such is not always evident, and seldom demonstrable. MCDM-based solutions do not necessarily optimize all of the objective functions under scrutiny. Consequently, the Pareto-optimal (or Pareto-efficient) solution concept has been introduced into multi-criteria modelling, and we have adapted it to our own system in its application to the case at hand. There usually exist a number of Pareto-optimal solutions; when specified, these may be considered as candidates for a final decision or resolution. A remaining issue, however, is how decision makers might decide a best resolution from a set of Pareto-optimal solutions (Korhonen and Wallenius, 1988; Korhonen, Wallenius, and Zionts, 1981).

2.1. A note on Pareto solutions. Pareto solutions are reliable because they resist extraneous influences – in other words, they are robust (Messac, 2000; Messac and Hattis, 1996; Messac and Chen, 2000; Messac and Sundararaj, 2000a; 2000b). Recent advances in design methodology attempt to incorporate robustness into design decisions: robustness analysis may be carried out by ascertaining whether a prescriptive decision changes when a datum, or data, or ensuing evaluation is removed from a modelling set. To ensure robustness in this regard, we constructed new data sets consisting of one observation less than the total in the original baseline, so that reliability could be sustained under modelling variations (Taguchi, 1993).

The advancement of robust design methods in statistics has focused on the improvement of the efficiency of Taguchi’s experimentation strategy and the modification of the signal-to-noise ratio as the pertinent design criterion (Box, 1988; Nair, 1992), for instance in nonlinear, programming-based, problem-solving methods (Chen, Allen, Mistree and Tsui, 1996; Parkinson, Sorensen and Pourhassan, 1993; Sundaesan, Ishii and Houser, 1993). A comprehensive review of robust optimization methods developed by the engineering design community is provided in Messac and Sundararaj (2000a; 2000b) and Zeleny (1973). We also note in this connection that a general robust design procedure was developed by Chen Allen, Mistree, and Tsui (1996) to improve the feasibility and reliability of design solutions (Cf. Doltsinis and Kang, 2004).

A common way to address trade-offs among multiple objectives in MCDM is known as the weighted-sum method, in which a single objective is specified so as to optimize the weighted sums of several objectives. However, using this method for multi-criteria optimization has inherent drawbacks, particularly its failure to satisfy assumptions of convexity and uneven distribution of data points; these drawbacks are discussed in Das and Dennis (1997). In consequence, Chen Allen, Mistree, and Tsui (1996) have applied a combination of multi-criteria mathematical programming (MCMP) methods and principles of decision analysis to address the multiple aims of robust design. A major element of their approach is the use of evaluative synthesis in fashioning a compounded compromise objective. Studies by Parkinson, Sorensen and Pourhassan (1993), Yu (1973), Zeleny (1973), and others have considered so-called Compromise Programming, whose basic idea is to reach a Pareto-ideal solution in which each individual attribute under consideration would achieve its optimum value. If attributes are in conflict, then the analyst would seek a solution closest to the ideal solution desired by decision makers (Chen, Allen, Mistree and Tsui, 1996).

Brunsson, Ellmerer, Schaupp, Trajanoski and Jobst (1998) developed and tested a rule-based MADM system to screen officer personnel records in the first phase of a board review procedure. In testing the system, an experiment involving mock officer personnel records demonstrated that the underlying method was successful at record selection for actual incumbents. To the same end, Drigas, Kouremenos,
Vrettaros and Kouremenos (2004) present fuzzy-logic techniques they have used to assess enterprise profile data in the evaluation of certain job candidates; their study combines three methods (SAW, TOPSIS, WP) in a multi-criteria mathematical programming (MCMP) technique in which any one of these methods can be used as the selection approach and the other two as validating methods. The solution is ideal if all three methods combined in sequence yield selections of the same set of candidates in the same order (Drigas et al., 2004). We have adapted this methodology in our own study.

2.2. Weighting techniques. While all job criteria specified in any particular instance are by definition important, some are necessarily more important than others. Numeric weights may be assigned to different criteria and so used to calculate performance scores, in identifying a leading candidate according to the decision-set criteria determined for the specific exercise. Greater weights indicate relatively higher importance in what becomes a rank-ordering of multiple selection criteria.

Several studies (Albayrak and Erensal, 2004; Iwamura and Lin, 1998; Labib, Williams and O’Connor, 1998; Lai, 1995; Vaidya and Kumar, 2006) have determined the global priority weights for different management decision alternatives aimed at the improvement of HR performance outcomes. A detailed review of various applications in varied settings is provided by Vaidya and Kumar (2006). In setting weights, \( w_1 \) is the weight of the most important attribute, \( w_2 \) is the weight of the second most important attribute and so on, down to \( w_m \), the weight assigned to the least important attribute. The two vital properties of greatest significance to these weights can be shown as follows:

\begin{itemize}
  \item 0 \leq w_i \leq 1 \\
  \sum_{i=1}^{n} w_i = 1.
\end{itemize}

2.3. Objective assignment of weights. The entropy approach to weighting (Hung and Chen, 2009) may be used to establish comparatively objective values in a decision system. An entropy value is a measure of the degree of uncertainty in a system, so that entropy weighting is intended to remove uncertainty, such as that introduced by human bias (Hwang and Yoon, 1987; Atkins and De Paula, 2006; Baierlein, 2003; Deng, Yeh and Willis, 2000).

Let \( k = \frac{1}{\ln n} \), where “\( \ln \)” denotes the natural logarithm to the base ‘e’. The entropy measure \( e_j \)

\begin{equation}
    e_j = -k \sum_{i=1}^{n} p_{ij} \ln p_{ij}, \quad j = 1, 2, \ldots, m.
\end{equation}

The degree of divergence \( d_j \), signifying the inherent contrast of focal attribute \( X_j \), also bound by 0 and 1 (Deng, Yeh and Willis, 2000; Zeleny, 1982), is defined as:

\begin{equation}
    d_j = 1 - e_j, \quad j = 1, 2, \ldots, m.
\end{equation}

The weights ranging between 0 and 1, derived from entropy from the degree of divergence for each attribute, are given by Deng, Yeh, and Willis (2000) as:

\begin{equation}
    w_j = \frac{d_j}{\sum_{k=1}^{n} d_k}, \quad j = 1, 2, \ldots, m.
\end{equation}

3. SToW

SToW shall denote SAW, TOPSIS and WP operations in this sequential order. In this study, simple additive weighting (SAW), weighted product and the techniques for order preference by similarity to ideal solution (TOPSIS) are used in MCDM to model a process of selecting a suitable applicant for a high-level appointment, as specified further in sections that follow.

3.1. Study setting. This study considers \( A_1, A_2, \ldots, A_n \), a set of \( n \) possible or alternative candidates competing for specified job opportunities, based on \( S_1, S_2, \ldots, S_m \), a set of \( m \) attributes (which could alternatively be considered decisional objectives or selection criteria). The environment of the selection process consists of entries \( x_{ij} \) that reflect the occurrence of attributes obtaining for candidate \( A_i \) under attribute \( S_j \) as follows. Consider the following matrix (Table 1):

\begin{equation}
\begin{array}{cccc}
S_1 & S_2 & \ldots & S_m \\
A_1 & x_{11} & x_{12} & \ldots & x_{1m} \\
A_2 & x_{21} & x_{22} & \ldots & x_{2m} \\
& \cdot & \cdot & \ldots & \cdot \\
& \cdot & \cdot & \ldots & \cdot \\
& \cdot & \cdot & \ldots & \cdot \\
A_n & x_{n1} & x_{n2} & \ldots & x_{nm} \\
\end{array}
\end{equation}
The decision matrix obtained from this setting is as follows:

\[
D = \begin{bmatrix}
    x_{11} & x_{12} & \ldots & x_{1m} \\
x_{21} & x_{22} & \ldots & x_{2m} \\
    \vdots & \vdots & \ddots & \vdots \\
x_{n1} & x_{n2} & \ldots & x_{nm}
\end{bmatrix}
\]

(6)

Let

\[
p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij}}, \quad i=1,2,\ldots,m; \quad j=1,2,\ldots,n
\]

(7)

Then define the \( n \times m \) performance matrix

\[
P = \begin{bmatrix}
p_{11} & p_{12} & \ldots & p_{1m} \\
p_{21} & p_{22} & \ldots & p_{2m} \\
    \vdots & \vdots & \ddots & \vdots \\
p_{n1} & p_{n2} & \ldots & p_{nm}
\end{bmatrix}
\]

(8)

3.2. Simple additive weighting. SAW is an authoritative MADM mathematical method used to perform pragmatic operations, applicable to many multivariate settings including modelling carried out in managerial contexts (Dawes, 1990). Attributes are assigned to some level of importance. Relative weights are assigned in direct proportion to the level of importance given for each attribute. To compare alternative candidates for employment in our case study, an evaluation score is calculated as follows:

Consider the \( m \)-vector of weights given by the column vector \( \omega = (w_1, w_2, \ldots, w_m)^T \). Define the transformation

\[
V_j = X(\omega p) = (v_1, v_2, \ldots, v_n)^T, \quad \text{a } n \text{-column vector, where:}
\]

\[
v_i = \sum_{j=1}^{m} w_j p_{ij}, \quad i = 1, 2, \ldots, n.
\]

(9)

These values are the SAW scores to be compared. They are arranged from largest to smallest, then used to rank-order the corresponding candidates and to select the leading candidate.

3.3. TOPSIS. TOPSIS, as described by Hwang and Yoon (1987), is a method for classical MCDM with the underlying logic of defining the ideal solution sought (or the positive ideal solution, PIS) and its converse, the negative ideal solution (NIS). The PIS maximizes the benefit criteria and minimizes the cost criteria, whereas, conversely, the NIS maximizes the cost criteria and minimizes the benefit criteria (Wang and Chang, 2007; Wang and Lee, 2007). Thus, the PIS consists of all of the best values attainable from extant decisional criteria, whereas the NIS consists of all the worst attainable values. The optimal alternative is positioned at the shortest distance from the PIS and the farthest distance from the NIS. The procedure used in this case now follows.

3.4. Formulation of the TOPSIS solution. Let \( J = [\text{benefit attributes}] \) and \( I = [\text{cost attributes}] \). Various sources (Steuer, 1989; Seo, Sakawa, 1988; Yoon and Hwang, 1995) present the MADM problem as:

\[
\text{optimize } \{f_j(x), f_2(x), \ldots, f_n(x)\} \quad \text{such that}
\]

\[
x \in X = \{x: g_h(x) \leq \gamma \leq \alpha ; \quad h = 1, 2, \ldots, k\}
\]

where

\[
f_j(x) = \text{benefit objective for maximization, } j \in J,
\]

\[
f_i(x) = \text{cost objective for minimization, } i \in I.
\]

(10)

Hwang and Liu (1993) provide reference points of PIS and NIS respectively as:

\[
f_t^* = \max_{x \in X} f(x) \quad \text{for all } j \in J,
\]

\[
= \min_{x \in X} f(x) \quad \text{for all } i \in I.
\]

(11)

and

\[
f_t^* = \min_{x \in X} f(x) \quad \text{for all } j \in J,
\]

\[
= \max_{x \in X} f(x) \quad \text{for all } i \in I.
\]

(12)

where \( t = 1, 2, \ldots, m \).

Further, Hwang, Lai and Liu (1993) define the PIS as the solution of an equation (10) consisting of the individual best feasible solutions given by \( f^* = (f_1^*, f_2^*, \ldots, f_m^*) \), and the NIS as the solution of equation (11), given by \( f^\wedge = (f_1^\wedge, f_2^\wedge, \ldots, f_m^\wedge) \).

TOPSIS procedure

Step 1: Define
\[ p_j = \frac{x_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}}, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n \]  
\[ e_i = -k \sum_{j=1}^{n} p_{ij} \ln p_{ij}, \quad j = 1, 2, \ldots, n \]  
\[ k = \frac{1}{\ln n}. \]  
\[ d_j = 1 - e_j, \quad j = 1, 2, \ldots, n \]  
\[ w_j = \frac{d_j}{\sum_{k=1}^{m} d_k}, \quad j = 1, 2, \ldots, n \]  

**Step 2:** In line with Yoon and Hwang (1995), define:

\[ r_y = \frac{x_{ij}}{\left(\sum_{k=1}^{m} \sum_{p=1}^{n} x_{kp}^2\right)^{1/2}}, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n \]  

**Step 3:** Calculate:

\[ v_{ij} = w_{ij} r_{ij}, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n \]  

**Step 4:** Identify the PIS and NIS. The PIS and NIS, denoted by \( A^* \) and \( A^\wedge \) respectively, obtain as follows.

The PIS is:

\[ A^* = \left( f_1^*, f_2^* \ldots, f_m^* \right). \]  

The NIS is:

\[ A^\wedge = \left( f_1^\wedge, f_2^\wedge \ldots, f_m^\wedge \right). \]  

These steps lead to calculation of distance (separation) measures.

**Step 5:** Calculate the separation measure. In this step the concept of the \( n \)-dimensional Euclidean distance is used to measure the separation distances of each alternative to the PIS and NIS. The corresponding formulae are as follows.

The PIS is given by:

\[ S_i = \left(\sum_{j=1}^{m} (v_{ij} - f_j^*)^2\right)^{1/2}, \quad i = 1, 2, \ldots, m. \]  

The NIS is given by:

\[ s_j = \left(\sum_{i=1}^{m} (v_{ij} - f_j^\wedge)^2\right)^{1/2}, \quad i = 1, 2, \ldots, m. \]  

**Step 6:** Calculate similarities to (separation from) the PIS; the similarity measure used is \( C^*_i \), which is defined by:

\[ C^*_i = \frac{S_i}{S_i + s_i}, \quad i = 1, 2, \ldots, m. \]  

With properties

\[ \diamond \quad 0 \leq C^*_i \leq 1 \]  
\[ \diamond \quad C^*_i = 0 \quad \text{when} \quad A_i = A^\wedge \]  
\[ \diamond \quad C^*_i = 1 \quad \text{when} \quad A_i = A^* \]  

**Step 7:** Rank-order preferences – define the preference order. Choose an alternative with maximum \( C^*_i \) or rank alternatives according to \( C^*_i \), in descending order.

3.4. **Weighted product.** In developing a merit score in evaluating a candidate using a WP, the starting point is to obtain a vector of values indicating scores associated with the attributes or criteria or objectives associated with a decision problem (Yoon and Hwang, 1995).

Consider the weights \( \omega = (\omega_1, \omega_2, \ldots, \omega_m)^T \); the \( i^{th} \) WP score is then defined by:

\[ p_i^w = \prod_{j=1}^{m} (x_{ij})^{\omega_j}, \quad i = 1, 2, \ldots, n. \]  

These values are the WP scores to be compared so as to select the leading candidate for the job.

4. **Data analysis**

4.1. **Data source.** Data were supplied by a Human Resources official seeking selection of a suitable applicant for a deputy director position in a district of Pretoria city. The data consisted of a matrix of values corresponding to marks awarded in accordance to various selection criteria during employment interviews, also including assessments of each candidate’s curriculum vitae and his or her verifiable credentials. Further specification of settings is withheld for privacy reasons in consideration of research ethics.

4.2. **Data format.** The data considered appear in a matrix format. The rows contained candidate identities (disguised for anonymity), while the columns defined the multiple criteria that were used in the comparison of candidates and ultimate selection of the best-suited candidate.

4.3. **Statistical packages.** SPSS and STATA were used to perform the statistical operations entailed in modelling, as previously described. The results obtained were consistent.
5. Results

As previously noted, a real HRM exercise was utilized in this study. The criteria used were:

- \( X_1 \): important qualification(s);
- \( X_2 \): relevant experience;
- \( X_3 \): self-expressed suitability for the position;
- \( X_4 \): capability to use existing policies to raise performance standards;
- \( X_5 \): understanding of roles required by the position;
- \( X_6 \): human relations competencies, principally regarding stakeholder relations;
- \( X_7 \): financial management skills/proficiency;
- \( X_8 \): self-sufficiency (independence and initiative).

Initial data for assigning fitness scores corresponding to the various selection criteria were generated from curriculum vitae reviews and communications with the references given by candidates. Each criterion was scored on a scale of 1 to 10, where ‘1’ is the worst possible performance fit and ‘10’ the best possible performance fit. The authors suggest that the outlier scores of 1 and 10 should be avoided as much as possible, since, logically, no candidate is likely to be either completely unsuited or perfectly suited to the position, at least on the basis of résumé reviews and reference checks.

The ranking based on scores obtained for the different candidates for the executive post in question were as follows:

Post: Deputy-director position
District: Undisclosed, Gauteng Province
Five candidates: \( C_1, C_2, \ldots, C_5 \)

Actual scores are provided in Table 4:

<table>
<thead>
<tr>
<th>Criteria</th>
<th>X_1</th>
<th>X_2</th>
<th>X_3</th>
<th>X_4</th>
<th>X_5</th>
<th>X_6</th>
<th>X_7</th>
<th>X_8</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_1</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>C_2</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>C_3</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>C_4</td>
<td>9</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>C_5</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

The decision matrix, therefore, is

\[
D = \begin{bmatrix}
5 & 8 & 4 & 4 & 6 & 6 & 5 & 3 \\
6 & 7 & 8 & 5 & 3 & 8 & 6 & 5 \\
4 & 8 & 5 & 7 & 9 & 8 & 8 & 2 \\
9 & 9 & 4 & 6 & 4 & 4 & 5 & 3 \\
8 & 8 & 7 & 5 & 2 & 6 & 5 & 2
\end{bmatrix}
\]

Performance matrix

The performance scores (goodness-of-fit calculations) are found using

\[
p_{ij} = \frac{x_{ij}}{\sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij}}
\]

Then the performance matrix is:

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.022</td>
<td>0.035</td>
<td>0.018</td>
<td>0.018</td>
<td>0.026</td>
<td>0.026</td>
<td>0.022</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>0.026</td>
<td>0.031</td>
<td>0.035</td>
<td>0.022</td>
<td>0.019</td>
<td>0.035</td>
<td>0.026</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td>0.018</td>
<td>0.035</td>
<td>0.022</td>
<td>0.031</td>
<td>0.040</td>
<td>0.035</td>
<td>0.035</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>0.040</td>
<td>0.040</td>
<td>0.018</td>
<td>0.028</td>
<td>0.018</td>
<td>0.018</td>
<td>0.022</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.035</td>
<td>0.031</td>
<td>0.022</td>
<td>0.009</td>
<td>0.026</td>
<td>0.022</td>
<td>0.009</td>
<td></td>
</tr>
</tbody>
</table>

Calculation of weights

Deng, Yeh and Willis (2000) propose the use of entropy (as previously defined) to derive attribute weights. Entropy depends on “facts” presented in the data, and the degree of divergence \( d_j \) is calculated from the various entropies. Weights are then calculated from the value \( d_j \). Hence, these weights are considered to be objective, as previously discussed, i.e., limiting or eliminating human bias and resulting uncertainty.

Entropy values determination

As previously considered, entropy is an uncertainty-reduction method for determining the attribute weights with relative objectivity. Entropy values

\[
e_j = -k \sum_{j=1}^{n} p_{ij} \ln p_{ij}
\]

are:

<table>
<thead>
<tr>
<th></th>
<th>( \ln p_{ij} )</th>
<th>( e_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3089</td>
<td>0.3089</td>
</tr>
<tr>
<td>2</td>
<td>0.3659</td>
<td>0.3659</td>
</tr>
<tr>
<td>3</td>
<td>0.2806</td>
<td>0.2806</td>
</tr>
<tr>
<td>4</td>
<td>0.2750</td>
<td>0.2750</td>
</tr>
<tr>
<td>5</td>
<td>0.2448</td>
<td>0.2448</td>
</tr>
<tr>
<td>6</td>
<td>0.3101</td>
<td>0.3101</td>
</tr>
<tr>
<td>7</td>
<td>0.2896</td>
<td>0.2896</td>
</tr>
<tr>
<td>8</td>
<td>0.1751</td>
<td>0.1751</td>
</tr>
</tbody>
</table>

Degrees of divergence

Degrees of divergence \( d_j = 1 - e_j \) are:

<table>
<thead>
<tr>
<th></th>
<th>( d_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.6911</td>
</tr>
<tr>
<td>2</td>
<td>0.8241</td>
</tr>
</tbody>
</table>
Table 8 (cont.). Degrees of divergence

<table>
<thead>
<tr>
<th>( j )</th>
<th>( d_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.7914</td>
</tr>
<tr>
<td>4</td>
<td>0.7250</td>
</tr>
<tr>
<td>5</td>
<td>0.7552</td>
</tr>
<tr>
<td>6</td>
<td>0.6899</td>
</tr>
<tr>
<td>7</td>
<td>0.7104</td>
</tr>
<tr>
<td>8</td>
<td>0.8249</td>
</tr>
</tbody>
</table>

Weights

The entropy derived weights \( w_j = \frac{d_j}{\sum d_k} \) are:

Table 9. Weights

<table>
<thead>
<tr>
<th>( j )</th>
<th>( w_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1202</td>
</tr>
<tr>
<td>2</td>
<td>0.1103</td>
</tr>
<tr>
<td>3</td>
<td>0.1251</td>
</tr>
<tr>
<td>4</td>
<td>0.1261</td>
</tr>
<tr>
<td>5</td>
<td>0.1313</td>
</tr>
<tr>
<td>6</td>
<td>0.1200</td>
</tr>
<tr>
<td>7</td>
<td>0.1236</td>
</tr>
<tr>
<td>8</td>
<td>0.1435</td>
</tr>
</tbody>
</table>

Calculation of SAW values

Step 1: Method used

The initial selection method used is the weighted average, deriving the SAW scores:

\[
 a_i^w = \sum_{j=1}^{m} w_j x_{ij} .
\]

Table 10. SAW scores

<table>
<thead>
<tr>
<th>( i )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_i^w )</td>
<td>5.044</td>
<td>5.937</td>
<td>6.288</td>
<td>5.385</td>
<td>5.237</td>
</tr>
</tbody>
</table>

The leading three candidates in order of merit were tentatively found to be C_3, C_2, and C_4, consistent with these scores.

Step 2. Validation: TOPSIS and WP

TOPSIS

TOPSIS steps require solutions for the PIS and NIS before the index can be calculated.

The best feasible individual solution for PIS is \( f^* = (9 \ 9 \ 8 \ 7 \ 9 \ 8 \ 5) \).

The worst feasible individual solution for NIS is \( f^\wedge = (4 \ 7 \ 4 \ 4 \ 2 \ 4 \ 5 \ 2) \).

Using the weights from Table 9, the distance measures from PIS are calculated using:

\[
 S_i = \left[ \sum_{j=1}^{n} w_j (f^*_j - x_{ij})^2 \right]^{\frac{1}{2}} .
\]

The PIS values obtained are:

Table 11. Candidates with their PIS scores

<table>
<thead>
<tr>
<th>( C_i )</th>
<th>7.790965</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_2 )</td>
<td>5.291768</td>
</tr>
<tr>
<td>( C_3 )</td>
<td>4.262969</td>
</tr>
<tr>
<td>( C_4 )</td>
<td>6.779535</td>
</tr>
<tr>
<td>( C_5 )</td>
<td>7.163008</td>
</tr>
</tbody>
</table>

Distance measures from NIS are based on:

\[
 s_i = \left[ \sum_{j=1}^{n} w_j (x_{ij} - f^*_j)^2 \right]^{\frac{1}{2}} .
\]

The NIS values obtained are:

Table 12. Candidates with their NIS scores

<table>
<thead>
<tr>
<th>( C_i )</th>
<th>3.199898</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_2 )</td>
<td>5.699095</td>
</tr>
<tr>
<td>( C_3 )</td>
<td>6.727894</td>
</tr>
<tr>
<td>( C_4 )</td>
<td>4.211328</td>
</tr>
<tr>
<td>( C_5 )</td>
<td>3.827855</td>
</tr>
</tbody>
</table>

Then the TOPSIS index is calculated from:

\[
 T_i = \frac{s_i}{S_i + s_i} .
\]

The values obtained are as follows:

Table 13. Candidates with their TOPSIS scores

<table>
<thead>
<tr>
<th>( C_i )</th>
<th>0.29114</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_2 )</td>
<td>0.51853</td>
</tr>
<tr>
<td>( C_3 )</td>
<td>0.61214</td>
</tr>
<tr>
<td>( C_4 )</td>
<td>0.38317</td>
</tr>
<tr>
<td>( C_5 )</td>
<td>0.34828</td>
</tr>
</tbody>
</table>

The leading three candidates in the order of merit: C_3, C_2, C_4.

WP

\[
 p_i^w = \prod_{j=1}^{m} (x_{ij})^{w_j} .
\]

Table 14. Candidates with their WP scores

<table>
<thead>
<tr>
<th>( i )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_i^w )</td>
<td>9.7474</td>
<td>9.9479</td>
<td>9.9556</td>
<td>9.7907</td>
<td>9.7020</td>
</tr>
</tbody>
</table>

The leading three candidates are therefore in the following merit order: C_3, C_2, C_4.
Remarks
For the three methods applications, the candidates’ merit-based ordering is exactly the same. There may be cases where differences occur across methods, and then additional selectors are required to either eliminate or keep candidates in contention, until only the final, best-suited candidate remains.

Step 3: Select candidate for appointment
In this instance, the leading candidate is consistently chosen by all three methodological applications, and by follow-on selection decisional criteria, to be C3.

6. Conclusion and recommendations

6.1. Order of preference. The SAW scores derived from the data give the following order of preference:
C3 C2 C4 C5 C1.
The WP scores derived from the data give the following order of preference:
C3 C2 C4 C1 C5.
The TOPSIS scores derived from the data give the following order of preference:
C3 C2 C4 C5 C1.
The top three candidates are ranked on merit in the order C3, C2, C4 according to each and all of the three methods. The excluded alternatives, C1 and C5, were not in the top three in any of these instances. Note that the ordering of C1 and C5 varied under WP, but that variation was not at all consequential in the ranking of the top three candidates. In the consolidation of the three methods, the rank order of candidates is as follows:
C3 C2 C4 C5 C1.
The selected candidate, in terms of SToW, is therefore C3.

6.2. Discussion. As indicated in the foregoing discussion, any one of these three methods can be taken as the initial means for selecting an applicant and the other two as validating methods. Application of the three methods consistently yielded the first three leading candidates in the same order – a validation test is that the top three candidates are consistently ranked in the same order of preference by all three methods. Hence, if one of these candidates is eliminated from the top three by one of these methods, that candidate is lacking in one or more qualities identified with the task and role structures of the contested job. If a candidate consistently falls out of the top three tier by these several methods, that candidate should probably be eliminated from consideration, and without delay.
The remaining two candidates would then be ranked anew as if they are the only ones available for the prospective appointment. The candidate to be appointed shall be the one who consistently outranks the other or others.

6.3. Discussion. As indicated in the foregoing discussion, any one of these three methods can be taken as the initial means for selecting an applicant and the other two as validating methods. Application of the three methods consistently yielded the first three leading candidates in the same order – a validation test is that the top three candidates are consistently ranked in the same order of preference by all three methods. Hence, if one of these candidates is eliminated from the top three by one of these methods, that candidate is lacking in one or more qualities identified with the task and role structures of the contested job. If a candidate consistently falls out of the top three tier by these several methods, that candidate should probably be eliminated from consideration, and without delay.

If all three candidates are placed by all the methods in the top three tier, the one to be appointed is the one who consistently leads the other two. However, this device may not guarantee an optimal selection, as occurs in the scenario that follows.

Alternate, hypothetical scenario
A situation can be posited where the leading candidate somehow fails to place consistently as such, as the leading choice. An accessible example involves a situation where the leading candidates are placed as follows:
SAW : C1 C2 C3
TOPSIS : C2 C3 C1
WP : C3 C1 C2
This is not a case where applications of the three methods serve to eliminate less suited applicants or to decide on the best-suited one. The next section considers such rare possibilities, in which the methods consistently determine the three leading candidates but fail to isolate the best-suited candidate.

Suggestion in case of a complete tie
This section discusses two methods for dealing with the possibility wherein the top three candidates do not lead to the final selection because of tied scores.

Possibility 1: Remove the criterion found to be of least importance
This possibility requires elimination of the criterion with the least entropy weight, as it has least importance by definition; this allows for keeping the original weights unchanged. Then calculate SAW, TOPSIS and WP with the remaining criteria for the top three candidates.

Possibility 2: Use ranks to derive weights
In this case a reflexive method of ranking is used to derive relative weights. The given criteria are ranked by the users (managers, decision makers) from best (ranked 1) to worst (ranked k). The respective weights are then calculated from the ranks using:

\[ w_j = \frac{i}{m} \sum_{j=1}^{n} \frac{1}{j}, \quad i = 2, ..., n. \]
Using the mathematical result on the denominator, the calculation of weights will be:

\[ w_i = \frac{2i}{m(m+1)/2}. \]  

(29)

Then calculate the SAW, TOPSIS and WP for the three candidates, rate them again, and re-assign rank, which is determined thereby on the basis of merit.

**Remarks**

The calculation of the measures of fit for the just-described ranking objective does not have to be limited to three top candidates; rather, it can be applied to all candidates preliminarily found to be suitable for the job. However, one may be best served by keeping previously eliminated candidates from reconsideration even if they possess some desirable traits lacking among the top three. While reopening the candidate pool may be advisable in some instances, in most others the utility of previously eliminated candidates is to winnow out the least desirable candidate among the top three. Put differently, this device should sharpen and facilitate the ultimate selection decision, not lead to indecision at this late stage.

In general, this study recommends empirical methods such as MCDM, SAW, TOPSIS, and WP for use in merit selection for employment. Incorporating entropy weighting reduces uncertainty and ambiguity, for instance as these relate to personal bias, producing relatively if not absolutely objective selective criteria.

**Future research**

Research is needed to explore and revise NIS and PIS such that their mathematical properties can make them independent and comparable selection criteria. It would also be worthwhile to identify other methods that use NIS and PIS in such a way as to lend greater consistency to selection processes. Extensive research synthesis could well inform the further development of the combined methodologies used in this study.

While the authors strove to reduce bias and uncertainty in this applied exercise, we remained aware that there are real threats to validity and consistency in real-world settings. There are many qualitative considerations that influence item-level ranking differences across candidates, especially since these judgments are usually made by managers and not researchers. These influences include, among many other factors, inconsistent perspectives among managers involved in job specification and selection, insufficient inter-rater reliability in the application of selection criteria, and decision makers’ difficulty in addressing close candidate assessments. We dwelt on this last difficulty in the closing alternative scenario: Here decision makers were bound to be stymied, if only because they lacked overarching and specifiable decisional criteria. Questions then suggest themselves for further development of our research: What are the tie-breakers to be in such instances of undecidable, close assessments? Are they to be prescribed competencies, role considerations, or other job factors or managerial demands? Should decision makers fall back on interviewing or testing results in instances like these?

It is well known that there is overreliance on interviewing in HR practice, and that behavioral interviews are notoriously unreliable (Barclay, 2001). The inadequacy of personality inventories and in-basket performance tests (and tests in general) is also widely acknowledged in the field (Barrick, Mount and Judge, 2001; Capraro and Capraro, 2002). These HRM standbys cannot plausibly serve to break an impasse when empirical methods fail to produce a clear standout candidate in a selection sequence. The question then becomes: What are the ultimate decisional criteria to be, and can they be adequately specified, so as to be effective when difficult choices are involved?

In proposing and using a MCDM system, the authors did not lose sight of the challenges involved in implementing complex decision criteria. These same challenges would indicate directions for future research: What are the truly relevant selection attributes in a given instance, and how are these to be selected? What are the relationships among selected attributes, and what is to be done when these conflict with one another? How do we structure attributes within a dynamic, evolving, and therefore changeable complex of values? How do we aggregate essential attributes in an overarching evaluation? In specifying utility functions, how do we map from lower level attribute item values to a higher level aggregate value set? Finally, can we provide for the often fungible quality of candidate qualifications. For example, how do we weigh experience against education, or substitute elements of one for another, especially if decision makers need to alter their relative weights when faced by actual candidates?

In proposing and testing empirical methods for employee selection, the authors never intended for these to substitute for quality judgments on the part of decision makers. In acknowledging the importance of judgment, however, we return to the original dilemma that drove the development and application of our methodological system. If
decision makers cannot sufficiently specify their criteria for a particular choice, and if they cannot control for human bias, where then may balance be found between decision systems and managerial judgment? These questions, once again, suggest research directions following on our study.

References