

DEVELOPMENT AND VALIDATION OF AN INSTRUMENT MEASURING THE STRENGTH OF THE HUMAN RESOURCE MANAGEMENT SYSTEM

DESENVOLVIMENTO E VALIDAÇÃO DE UM NOVO INSTRUMENTO PARA MEDIR A FORÇA DO SISTEMA DE GESTÃO DE RECURSOS HUMANOS

*Jorge F. S. Gomes
Joaquim P. Coelho
Anabela Correia
Rita C. Cunha*

ABSTRACT

Notwithstanding the theoretical and empirical support for the human resource management-organisational performance connection, authors still do not know *how* this association works. Bowen and Ostroff (2004) introduced the concept of the strength of the human resource system. In such systems, messages regarding what is appropriate behaviour are sent to employees in an unambiguous and consensual way. Human resource strength affects the way people interpret their environment (situation strength): strong human resource management leads to stronger situations, whereas weak human resource management leads to weaker situations. This research presents an instrument aimed at measuring the concept of strength; furthermore, it assesses Bowen and Ostroff's (2004) model.

Data was collected via questionnaires in six companies, from two distinct studies. Study I is based on 90 questionnaires from five companies; results from this study allowed us to improve the quality of the instrument developed, which in turn was used in a second study, carried out in a single company (88 valid questionnaires).

Overall, results show good reliability estimates of the new instrument, as well as a partial confirmation of Bowen and Ostroff's (2004) model. Explanations for these mixed results are presented and discussed. These explanations may be related to the research design itself, but they may also be due to unsatisfactory or inadequate definition in some of the elements of Bowen and Ostroff's model. Implications for future research are examined.

Keywords: Human resource management, instrument development, strength.

RESUMO

O conceito de “força” (Bowen & Ostroff, 2004; Schneider, Salvaggio & Subirats, 2002; Shein, 1981), nas ciências organizacionais e gestão, é dotado de um elevado potencial teórico, mas ao mesmo tempo é intrigante e mágico. O potencial teórico deriva do facto que “força” reflecte o cruzamento de dois ou mais níveis de análise, pelo que pressupõe a ligação das percepções individuais aos comportamentos do grupo e da organização. O enfoque em múltiplos níveis de análise há muito que representa um dos grandes desafios em investigação e intervenção nestas áreas. O carácter intrigante e mágico reflecte a ideia de que o conceito de força capta e reproduz o pensamento e a crença de colectividades humanas. Existe ainda uma outra suposição: a de que aquilo que o grupo é capaz de fazer, é superior ao que o

indivíduo isolado consegue realizar. Neste sentido, “força” é a sabedoria das multidões, para usar um título de um livro por Surowiecki (2005).

Nesta apresentação para o CIEO, o conceito de força da GRH (Bowen & Ostroff, 2004) é explorado num modelo em que se procura explicar o desempenho organizacional mediado pela força do clima. Ademais, propõe-se que a liderança concorre com a força da GRH para influenciar o desempenho através do clima.

O trabalho empírico decompõe-se em duas fases. Numa primeira, foram realizados vários estudos com vista ao desenvolvimento e teste de uma nova medida de força da GRH. Na segunda fase, utilizou-se um questionário num grande grupo hoteleiro, para recolher mais de 500 respostas válidas, e com o qual se procurou medir os conceitos centrais no modelo. Na sessão para o CIEO apresentam-se os resultados principais do estudo da 2ª fase, e mencionam-se brevemente os resultados da 1ª fase. Sugerem-se ainda pistas para interpretar os resultados, assim como se avançam ideias para a continuação do projecto.

Palavras-chave: Gestão de recursos humanos; desenvolvimento instrumental; força

JEL Classification: M12

1. INTRODUCTION

“How much does human resource management matter?” This is the question with which Gerhart, Wright and McMahan open their year 2000 article published in *Personnel Psychology*. The question reflects the debate in human resource management (hereinafter HR) literature in recent years, as the personnel function is increasingly called upon to show how it can contribute towards increasing individual and organisational productivity, in particular, and towards supporting and reinforcing organisational goals and missions, in general. These concerns have been at the centre of an important stream of research in HR in the last few decades, known as the strategic HR (SHR) perspective. Despite such interest, some studies (e.g. Aijala, Walsh and Schwartz, 2007) show that the HR function still has a long way to go before it reaches alignment on strategic priorities.

SHR focuses on the ways in which HR practices and the HR system are critical to organisational effectiveness. The implicit assumption is that if HR is done well, this will somehow make organisations perform more effectively (Ferris et al., 1998). The key research issue is captured by the word *somehow* in Ferris and his colleagues’ assertion. In fact, notwithstanding the theoretical and empirical support for the HR-organisational performance relationship, there is still a significant lack of knowledge with regards to *how* such association works.

Following this plea, Bowen and Ostroff (2004) proposed a model in which the relationship between HR and performance is mediated by the psychological interpretation of events, and specifically by the extent to which people diverge or converge in their views of situations. The degree of convergence is called psychological strength, and it reflects the degree of harmony that exists among the organisation’s collaborators, as far as their perceptions and beliefs are concerned. Greater harmony indicates stronger psychological situations, whereas greater disharmony denotes weaker psychological situations.

The main influencing cause of psychological strength, according to Bowen and Ostroff (2004), is a novel concept: the strength of the HR System. How strong an HR system is depends upon a set of attributes of such a system, which transcends any existing HR

practices, policies and philosophies. These meta-attributes are related more to the way the HR communicates with employees than to the content of the messages it conveys to people in the organisation. In other words, such features tell us about *how* HR communicates, not about *what* it communicates.

Bowen and Ostroff's (2004) model is a refreshing step in a promising stream of research in HR. However, it has not been tested for its conceptual validity or its power of explanation. The aims of the current paper are twofold: 1) to advance an operational definition of the concept of HR strength; and b) to present an initial test of Bowen and Ostroff's (2004) model.

2. LITERATURE REVIEW

2.1 The strength of human resource systems

Mischel (1973) first proposed the concept of situation strength to explain the influence that situations have on shaping individuals' behaviours. According to the author, in certain situations, behaviour is the result more of people's inner states than anything else, whereas at other times, it is the situation that shapes the actions of individuals. He goes on to elucidate that a situation is called 'strong' when it is able to lead everyone to construe particular events in the same way, and it induces uniform expectancies regarding the most appropriate response pattern. Conversely, individual differences determine behaviour when the situation is ambiguously structured and people have no clear expectations about the behaviours they need to adopt; these are called 'weak' situations. In sum, situational strength deals with the extent to which a situation induces conformity – a strong situation – or is interpreted as ambiguous – a weak situation.

Bowen and Ostroff (2004) picked up these ideas, further proposing that in strong psychological situations people share interpretations of particular events, therefore behaving in a uniform and consistent way, and showing a similar orientation towards goals and desired standards of performance. In weak situations, people rely on their predisposition states in order to behave, and it is therefore likely that they will exhibit more differences in terms of what is acceptable and efficient organisational behaviour.

Situation strength is not a novel notion. In fact, there have been some developments of this concept in the culture and climate literature. Climate strength is the extent to which people share perceptions regarding various issues related to their environment, i.e. the organisation (Schneider, Salvaggio, and Subirats, 2002; Payne, 2000). Culture strength (Schein, 1981) refers to the degree to which people share values and beliefs regarding their work and their organisation. Thus, both climate strength and culture strength can be used as proxy constructs to evaluate situation strength.

A novel idea introduced by Bowen and Ostroff (2004) is that of the strength of HR systems. These systems share three characteristics: a) distinctiveness (when a particular situation stands out in the environment, thereby capturing attention and arousing interest); b) consistency (it refers to an HR function which communicates regular and consistent messages over time, people and contexts); and c) consensus (agreement among employees in their view of the event-effect relationship).

These three characteristics were originally put forward by Kelley (1973), building on Heider's theory of attribution. According to Kelley (1967), attribution is an inferential process that allows individuals to understand and to be acquainted with the external world, mainly through objective information processing. When explaining behaviours in their environment, people take into account information about three distinct sources: objects in the environment (called entities), persons interacting with these objects, and the consistency

of these behaviours (circumstances). The output of this process is an attribution of the effect to an external cause (or to the entity itself) or to the self (internal attribution). Whether one makes attributions to internal or external causes depends on how distinct, consistent and consensual the situation is. These three main criteria help people confer validity to their knowledge about the external world. To the extent that one's attribution fulfils these criteria, there is high confidence in attributing the effect to the entity, that is, to reality. When one's attribution does not satisfy these criteria, judgement becomes uncertain and attribution to the entity is replaced by attribution to one's personal characteristics.

Back to Bowen and Ostroff's (2004) assertions, in strong HR systems, messages regarding what is appropriate behaviour are communicated (via HR practices) to employees in an unequivocal, consistent and consensual way. Hence, HR systems must possess a set of unique characteristics, which are related to the *process* by which a consistent message about HR *content* is sent to employees. The authors define the three dimensions in terms of nine attributes:

- Distinctiveness, which has four dimensions:
 - Visibility or salience: this refers to the degree to which HR practices are salient and readily observable. In social psychology, salience of an object is important so that people are able to make sense of it. For example, if performance criteria are not transparent, this will not create Mischel's strong situation;
 - Understandability: lack of ambiguity and ease of comprehension of HR practice content. HR practices (situational stimulus) can be ambiguous or understandable; in the first case, people will have different interpretations of a particular practice, whereas in the second case they will probably share interpretations;
 - Legitimacy of authority: degree to which the HR system and its agents have legitimacy in moulding people's and groups' actions. If the HR function is perceived as a high-status and high-credibility activity, then it is likely that individuals perceive it as an authority situation, thus submitting themselves to performance expectations as formally sanctioned behaviours; for example, if the HR director sits on the company's board, people may interpret it as "people matter in my company";
 - Relevance: degree to which people perceive the HR system as an important factor in helping to achieve particular goals in the company. The link between individual and organisational goals is central here: individuals must perceive that organisational goals are an important means to achieving their personal goals, and vice-versa.

The above features help draw attention to the message conveyed by the communicator (the HR system), thereby increasing the probability that the HR message will be interpreted uniformly among employees. But this does not suffice, since people will need to perceive that the HR function behaves in a consistent and regular way in all situations, with all employees, and over time. This is the second characteristic:

- Consistency, which refers to an HR function that communicates regular and consistent messages over time, people and contexts. It has three dimensions:
 - Instrumentality: this concept is similar to Vroom's instrumentality concept (Yukl, 1998) in his expectancy model of motivation. It refers to the extent to which people perceive that the HR function and its communicators link outcomes to behaviours or performances in a timely and consistent manner;
 - Validity: to what extent the purpose of HR practices is congruent with what is actually done in practice; for example, if "innovative and creative behaviour" is a key criteria in performance appraisal, but in practice nobody cares about someone's ideas and suggestions, then there is a low perceived validity in the performance appraisal;
 - Consistent HR messages: compatibility and stability of the signals sent by the

HR practices. Consistency has three dimensions: i) what senior managers say are the organisation's goals and values, and what employees actually conclude those goals and values to be; ii) consistency across HR practices; and iii) consistency over time.

- Consensus: agreement among employees in their view of the event-effect relationship. Several factors can help foster consensus among employees and can influence whether individuals perceive the same effect with respect to the entity or situation in question. Among these are:
 - Agreement among principal message senders: degree to which key decision-makers (Bowen and Ostroff write about two: top managers and HR executives) are seen as agreeing with each other in respect to the message sent by the HR function;
 - Fairness of the HR system: to what extent employees perceive that the HR system complies with the three dimensions of justice: procedural, distributive and interactional.

2.2 The measurement of HR strength

Bowen and Ostroff's framework is appealing and it offers a provocative set of ideas. However, it needs to be tested and refined, so that further investigation can be carried out along these lines. The first question driving the current research is, therefore, "how can these nine attributes be assessed?"

Studies measuring the concept of the strength of HR Systems are still scarce in the literature and most are not concerned with the development of a reliable and valid way to measure the concept. In this section, a brief overview of existing studies is provided.

Edgar and Geare (2005) developed 20 questions to measure the perceived strength of four HRM practices (health and safety, training and development, equal employment opportunity, and recruitment and selection), which was operationally defined as the degree to which people agreed or disagreed with each statement (e.g. "working conditions are good"). Since the authors did not aim to explore Bowen and Ostroff's meta-attributes, no other information is provided in their work.

Dorenbosch, Reuver and Sanders (2006) carried out research in 66 hospital departments from four hospitals, in order to study the consensus between line managers and HR professionals. The Dutch team limited their investigation to two of the nine attributes in the Bowen and Ostroff (2004) model: consensus on the HR message and legitimacy of the HR message. Their aim was to check the effect of these two features on commitment strength, defined as the collective degree of identification and dedication to the organisation and organisational goals. A questionnaire with a similar set of questions was used to assess the views of line managers and HR professionals as far as two issues were concerned: human resource practices and the HR function roles. Agreement among the two groups of message senders (line managers and HR professionals) was calculated as the absolute deviance scores of the mean scores on the HR practices. The authors then calculated the inverse deviance scores, so that high scores on consensus refer to high agreement on the HR message among line managers and HR professionals. Legitimacy of the HR message was also assessed by means of a questionnaire, namely by asking line managers and HR professionals about their perception of the role of the HR function. The four roles put forward by Ulrich (1997) were: strategic, partner, change agent, administrative expert, and employee champion.

Finally, Chen, Lin, Lu, and Tsao (2007) collected data through questionnaires from over 400 hairdressers and shop-owners, to study the relationships between employee affective commitment, employee perceptions of HR practices and job performance. The Taiwanese team used HR Strength as a moderating variable in their model, and they focused solely on one of the attributes in Bowen and Ostroff's (2004) model: consistent HR messages between

hairdressers and shop-owners, as far as five HR practices were concerned: employment security, recruitment and selection, training, employee status, and compensation. The reason why the researchers used consistent HR messages as a proxy for the whole HR Strength concept is that they based their operational definition on a previous work by Ostroff and Bowen (2000), in which they proposed that HR Strength can be determined by how closely employee perceptions match those of their managers.

2.3 Issues in measuring HR strength

From the above discussion, there are a number of key elements which are important when using the concept of HR Strength to address research in the human resource field.

1. Firstly, one needs to acknowledge that the nine aforementioned attributes are grounded on the concept of HR practices, which is a lower-order concept, in contrast to the concepts put forward by Bowen and Ostroff, which are higher-order (more abstract) ones. This implies that each of the attributes needs to be assessed by reference to lower-level concepts. Hence, people need to think over an HR practice or a set of interrelated HR practices (an HR System; Delery and Doty, 1996), and only then can they judge if that particular HR practice or HR System is distinct, consistent and consensual.
2. Secondly, as argued by Bowen and Ostroff (2004), the concept of strength requires the judgements and perceptions of employees; therefore the best way to assess the meta-attributes is by having the employees themselves make the assessment. The appropriate unit of measurement is the individual. However, employees do not always have enough information regarding a particular HR practice or set of practices. For instance, most employees would not have enough information regarding how recruitment and selection is carried out in their organisation, since they may have had only one or two contacts with such practices (i.e. when entering their organisation). Therefore, they need to be called upon to evaluate practices with which they have regular contact, such as training or performance appraisal.
3. The concept of strength is based on what James (1982) and Chan (1998) called compositional models, i.e. constructs operationalised at one level of analysis which are then somehow represented at another level of analysis. These compositional models allow good multilevel analysis to be carried out (Chan, 1998). The implication of this literature is that in order to operationalise HR Strength, one should first create a measure by which people are called upon to evaluate some feature of HR (e.g. a practice), and then the construct of "strength" is derived e.g. from some dispersion measure. In addition to dispersion and other similar measures, the more "traditional" measures of central tendency also allow interesting relationships to be studied, as recognised by Bowen and Ostroff (2004). Other issues in multilevel research are pointed out by authors such as James, Demaree, and Wolf (1984), Podsakoff and Organ (1986), and Podsakoff, MacKenzie, Lee, and Podsakoff (2003), and they include: hetero and auto-reported measures, data collected from multiple sources of information, and group and organisational indicators built out of information collected at the individual level.
4. No easy answer can be found in the literature as to which HR practices should be selected. At least two interpretations are found: a) HR practice as a feature of a particular HR activity (e.g. recruitment is the activity, and the feature is internal recruitment, external recruitment, or both); or b) HR practice as the set of specific activities aimed at achieving a particular objective or group of objectives (e.g. HR practices directed at supporting innovation). The first view allows comparison across industries and companies, since it deals with generic HR practices which supposedly can be found everywhere. People can also produce meaning regarding the overall HR generic orientation, HR goals, or HR roles, as shown in some works (e.g. Arthur, 1994; Dorenbosch et al., 2006; Ulrich, 1997; Ulrich & Brockbank, 2005), but this sensemaking activity (Weick, 1995) is pretty much dependent on more visible activities such as training and performance appraisal. As a re-

- sult, the current work focuses on generic HR practices that can be found in any company.
5. Finally, a related question is which practice or set of practices should be used to collect people's interpretations? The literature is not conclusive in this regard. For example, Pfeffer (1994) argues for a greater use of 16 practices, such as selectivity in recruitment and high wages. Ichniowski, Shaw and Prensushi (1997) used 8 practices in their study, including communication and labour relations. Dorenbosch et al. (2006) used 7 practices, such as appraisal outcomes and appraisal criteria. Chen and colleagues (2007) used 5 practices. Combs, Hall and Ketchen (2006) estimate that, on average, studies use seven practices when conducting research on the HR system.

2.4 Research goal

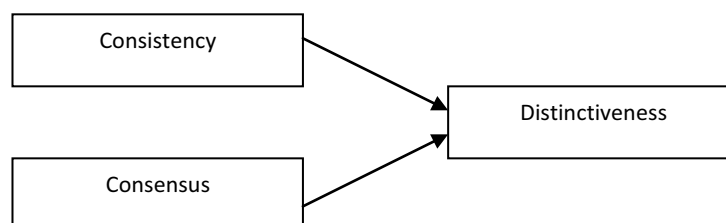
Bowen and Ostroff's (2004) model offers a refreshing and promising framework to look deeper into the relationships between the HR System, strategy, and performance. It links several levels of analysis in the organisation, and allows a dynamical view over process, human and management subject matters.

How each of the nine attributes relates to each other is still largely unknown. The authors suggest a few associations in their work, but in fact none of them has been truly tested so far, nor is their text dedicated to delving deeper into the matter. For this reason, as far as associations between the nine attributes were concerned, the current research followed an exploratory approach.

However, with regard to the three dimensions originally taken from Kelley's (1967, 1973) theory, a few associations can be drawn, which in turn require a confirmatory type of research. Although the three dimensions of distinctiveness, consistency and consensus are all required to help people build interpretations and make attributions about phenomena, the literature seems to defend the primacy of distinctiveness. Hewstone and Jaspars (1988), for example, conducted two studies in which consensus and consistency were found to influence distinctiveness, which in turn was the last and most powerful influencing factor in shaping final attributions. Bowen and Ostroff (2004) accept these relationships, further proposing that distinctiveness has the last word in influencing the Strength of the Situation.

Overall, these findings and theoretical propositions recommend that the three main dimensions studied in the current research should be related, as shown in figure 1.

Figure 1 – Model proposed for testing the three dimensions of HR System Strength



Since the whole model is dependent on a new concept, namely that of HR Strength, it is vital that researchers develop powerful and suitable measures to assess such important, and yet still questionable connections in the HR field. The purpose of the current investigation is to advance knowledge in this respect, by reporting the preliminary results of a research project on the effects of HR System Strength on individual and organisational performance. In particular, the current text presents and assesses a novel measure of HR Strength, tested in two studies in different settings.

3. METHOD

3.1 Participants

A research project was set to develop appropriate measures for testing the full conceptual model presented by Bowen and Ostroff, as well as to explore previous propositions in the HR literature in which the HR System is linked with strategy and individual and organisational performance through situation strength. The current text presents results from two studies aimed at developing and testing operational measures for the constructs in Bowen and Ostroff's model.

Both studies used a quantitative approach (questionnaire). All items composing the main scales required respondents to rate the degree to which they disagreed or agreed with the particular statement. A 1-7 Likert type of scale was used. This was a similar procedure to the one followed by Edgar and Geare (2005), i.e. 1 ("totally disagree") represents a weaker HR practice or goal, and 7 ("totally agree") represents a stronger HR practice or goal. Biographical data included gender, age groups, tenure, and function in the company.

91 people participated in the first study, from five companies in distinct economic sectors: (company 1 – 16.5%), construction (company 2 – 19.8%), maritime management (company 3 – 20.9%), commercial printing (company 4 – 16.5%), and elevators (company 5 – 26.4%). 62% of respondents are male, and 26% are female (2.2% missing). 27.5% of respondents fall into the 30-40 age category.

Coincidentally, the second study is also based on 91 people, though from a single, large telecommunications company. We decided to focus on one single company in order to eliminate any company effects, potentially active in study 1. One such effect was the professional activity performed by respondents. While in the first study there were as many as 30 different functions across the five companies, in the second study, 86% of the respondents fall within the same job category, namely call centre operators. 42% of respondents in study 2 are male, and 52% are female (6.6% missing). 19.8% of respondents fall into the 30-40 age category.

3.2 Item development procedures

Due to the exploratory nature of the current work, and after a set of five interviews with HR managers, the authors decided to use performance appraisal in study 1, as the framework on which items could be generated to assess the nine attributes proposed by Bowen and Ostroff (2004). In fact, at the time of the interviews, performance appraisal was an important goal in the participating companies, with processes being implemented or already in place. In study 2, carried out shortly after the first one, it was decided that the main focus would be the HR System as a whole (e.g. generic orientation and HR goals), without emphasising any particular HR practice. The change is justified since the authors wished to engender a feeling as to which choice would work better: to focus on one single HR practice or to focus on the HR function. Other differences introduced from study 1 to study 2 are explained below.

The researchers used an iterative process to generate suitable sentences to compose the questionnaire, following authors such as Nunnally and Bernstein (1996). A first set of 63 sentences covering all the nine attributes was first created by the authors of this research and some fellow HR professionals (three organisational behaviour master's students at the time of the study). The phrases were then mixed up so that no association between the indicators and the constructs was possible to identify. The 63 questions were then circulated within the group, for blind recoding into the original nine constructs.

A final stage involved computing an inter-coder agreement index (percentages of matches) for all 63 sentences. Only sentences with three matches (50%) or above were selected for inclusion in the questionnaire. There were 44 sentences in this situation. Further

refinement led to the elimination of a few sentences which were redundant, somehow confusing, inadequate, or had only 50% of matches. In the end, there were 36 sentences in the questionnaire, four per attribute. Table 1 shows an example of the sentences generated with this procedure.

Table 1 – Items used in the HR Strength scale – selected examples from study 1 (before change

Attribute	Examples of indicators
Visibility	“In this company, performance appraisal goals are known to everyone”
Instrumentality	“Someone who is a bad performer in this company, should expect to feel the consequences (e.g., no yearly rewards)”
Agreement among message senders	“HR Managers in this company share the same vision as Senior Managers with regard to performance appraisal principles”

In the second study, items were changed in order to accommodate the shift in content. However, following preliminary analysis on the data from study 1, other improvements were introduced. In particular, and since several sentences did not render sufficiently good results in study 1, more sentences were created in study 2, while others were eliminated.

75 new sentences were written or rewritten and submitted to validation content by nine HR professionals. Based on their comments, some sentences were reworded. The six most consensual items of each attribute were chosen to integrate this version of the questionnaire. For “Fairness of the HR system”, the authors used a scale already adapted to a Portuguese context, based on the works of Rego (2000, 2002). Altogether, there were 54 items in the final version of the questionnaire used in study 2.

Table 2 shows an example of the sentences from the questionnaire used in study 2. As can be observed in tables 1 and 2, changes were sometimes very small, while in other cases, they required a completely new sentence.

Table 2 – Items used in the HR Strength scale – selected examples from study 2 (before changes)

Attribute	Examples of indicators
Visibility	“HR practices are known to everybody in this company”
Instrumentality	“If my behaviours in this company are adapted to its culture, I know that I’ll be rewarded in some way”
Agreement among message senders	“HR Managers in this company share the same vision as Senior Managers”

In both studies, the questionnaires were subject to thorough qualitative testing before application. Several HR managers, HR professionals and the research team revised each sentence for its clarity, face validity, and simplicity.

3.3 Variables

The nine attributes put forward by Bowen and Ostroff (2004) were the main target in both studies. In study 1, there were 4 sentences per attribute, whereas in study 2, there were 6 sentences per attribute. Since this framework was based on Kelley’s (1967, 1973)

attribution theory, the three dimensions of distinctiveness, consistency, and consensus were also included in the testing procedures. For terminology reasons, hereinafter the word ‘dimensions’ refers to Kelley’s (1967, 1973) three concepts; the words ‘attributes’ or ‘meta-attributes’ are used to refer to Bowen and Ostroff’s (2004) nine concepts.

Figure 2 shows these two levels of variables.

Figure 2 – Variables analysed in the current research

Distinctiveness	Visibility or salience Understandability Legitimacy of authority Relevance
Consistency	Instrumentality Validity Consistent HR messages
Consensus	Agreement among message senders Fairness of the HR system

4. RESULTS

4.1 Plan of analysis

Before the main data analysis was run, data screening was performed using SPSS 17. A first set of statistical procedures aimed at exploring the data. This had several purposes: to verify the data’s accuracy, to identify missing values, to spot outliers, and to test assumptions of multivariate statistical techniques. This was done for all levels of variables. Overall, only one case was removed from study 1 due to excessive missing data and 3 outliers were deleted from study 2, since they registered p values smaller than 0.001 on the Mahalanobis d-squared’ test. Final sample sizes were: $n_1=90$ and $n_2 = 88$.

Most individual variables showed normal or approximately normal distributions, and the same was true for the aggregated constructs.

In a second phase, hypothesised models were analysed by computing structural equation models (SEM) using AMOS 17 (Arbuckle, 2007). The maximum likelihood method (MLM) was used for parameter estimation, which is taken as a robust method. The models’ goodness-of-fit was assessed using several indices. The absolute goodness-of-fit indices computed were the chi-square goodness-of-fit statistic (CMIN) and the Root Mean Square Error of Approximation (RMSEA). The relative goodness-of-fit indices were: the normed chi-square (CMINDF) and the Comparative Fit Index (CFI). For model comparison, the Browne-Cudeck Criterion (BCC) and the Bayes Information Criterion (BIC) were used. BIC has a greater tendency to pick parsimonious models than BCC. Bootstrapping was also used for several estimation and comparison proposals.

4.2 Sensibility and reliability estimates

Table 3 shows means and standard deviations for all the variables (original scales ranged between 1 and 7). All items and latent variables show approximately normal distribution, checked by Kolmogorov-Smirnov test and/or the skewness and the kurtosis coefficients.

Table 3 – Descriptive statistics for studies 1 and 2

Variables	Study 1 (n=90)	Study 2 (n=88)
Visibility or salience	4.23 (1.17)	4.51 (0.97)
Understandability	4.20 (1.14)	4.63 (1.04)
Legitimacy of authority	4.33 (0.95)	4.65 (1.13)
Relevance	4.15 (0.98)	4.61 (1.18)
Distinctiveness	4.22 (0.98)	4.61 (1.11)
Instrumentality	4.29 (1.25)	4.81 (0.98)
Validity	4.31 (1.05)	4.78 (0.96)
Consistent HR messages	4.09 (0.99)	4.58 (1.06)
Consistency	4.23 (1.01)	4.69 (1.16)
Agreement among message senders	4.36 (1.02)	4.48 (1.02)
Fairness of the HR system	3.96 (1.25)	4.25 (1.07)
Consensus	4.16 (1.06)	4.09 (1.08)

A second set of procedures intended to investigate reliability, which was computed through alpha coefficients. Table 4 shows results for both studies.

Table 4 – Reliability (Cronbach’s alpha) estimates for studies 1 and 2

Variables	Study 1 (n=90)	Study 2 (n=88)
Visibility or salience	0.86	0.75
Understandability	0.86	0.80
Legitimacy of authority	0.79	0.89
Relevance	0.82	0.89
Distinctiveness	0.95	0.94
Instrumentality	0.85	0.75
Validity	0.82	0.85
Consistent HR messages	0.79	0.86
Consistency	0.93	0.92
Agreement among message senders	0.75	0.83
Fairness of the HR system	0.91	0.70
Consensus	0.90	0.83

The first two sets of procedures suggest that the item-generating procedures were efficient insofar as they produced sentences which were capable of discriminating between respondents, and also that they maintained a good level of internal consistency. Overall, results do not differ sharply between the two studies, although ‘fairness of the HR system’ registered a fall in reliability from study 1 to study 2.

Distinctiveness

To assess the quality of dimension distinctiveness, principal components analysis (PCA) was used in both studies.

In study 1, two components were obtained, explaining 65.2% of total variance. The components were rotated using iterative varimax method. Items with high loadings on component 1 were originally designed to measure ‘visibility’ and ‘understandability’; component 2 is mainly composed of items which were firstly designed to measure ‘legitimacy’ and ‘relevance’.

Confirmatory factor analysis (CFA) was next used with the items which revealed weights higher than 0.5 in each component and also that did not load on both components. These two criteria meant that 11 items were accepted for CFA (3 for ‘visibility’ and 3 for ‘understandability’ in factor 1; 2 for ‘legitimacy’ and 3 for ‘relevance’ in factor 2). Results from CFA confirmed the existence of two factors in dimension distinctiveness.

After some further analysis, parsimonious and good model fit was achieved with 8 items (2 for each of the 4 attributes considered earlier). For this last model, the goodness-of-fit indicators were: CMIN=25.507, DF=19, CMINDF=1.342, CFI=0.985, RMSEA=0.062, BCC=63.332 and BIC=102.004.

In study 2, four components were obtained, explaining 73.3% of total variance. However, this exploratory analysis was deemed inappropriate since attributes showed high correlations amongst them. On the other hand, CFA indicated a good fit for 3 subscales (‘visibility’ and ‘understandability’ together, ‘legitimacy’, and ‘relevance’). There were also 11 items retained after this analysis. For this model, the goodness-of-fit indicators were: CMIN= 47,302, DF=42, CNINDF= 1,126, CFI=0,990, RMSEA=0,038, BCC=102,982 and BIC=154,758.

Consistency and Consensus

For latent variable consistency and consensus, both EFA and CFA were unable to support the hypothesis of more than one attribute per variable. In other words, it was not possible to clearly identify the attributes of ‘instrumentality’, ‘validity’, and ‘consistent HR messages’ in dimension consistency. Neither was it possible to find evidence in support of separate attributes ‘agreement among message sender’ and ‘fairness of the HR system’ in dimension consensus. This lack of discriminant validity is probably due to the high correlations among the attributes proposed by Bowen and Ostroff (2004).

In study 1, after a set of intercorrelations and reliability analyses, 10 items were considered for assessing consistency, and 5 items for measuring consensus. In study 2, 9 items were considered for consistency, and 6 for consensus.

HR System

A final set of procedures aimed to test the full model in each study. Figures 3a and 3b depict the final solutions; figure 3a shows results from study 1 and figure 3b shows results from study 2. Standardised weights for each path and squared multiple correlations for each dependent variable are represented by the numbers next to each path and/or variable.

Figure 3a – HR System: SEM Results from study 1

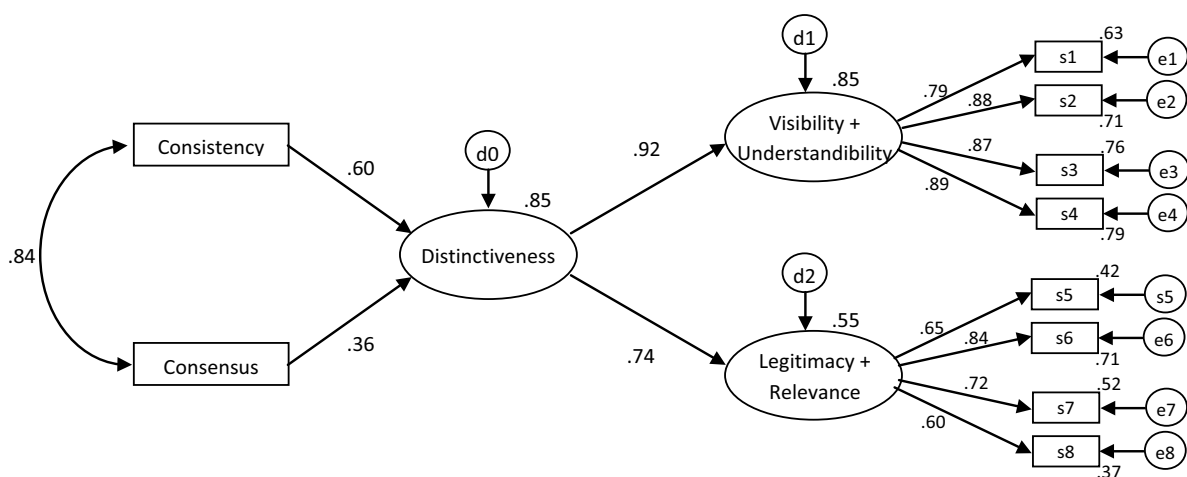


Figure 3b – HR System: SEM Results from study 2

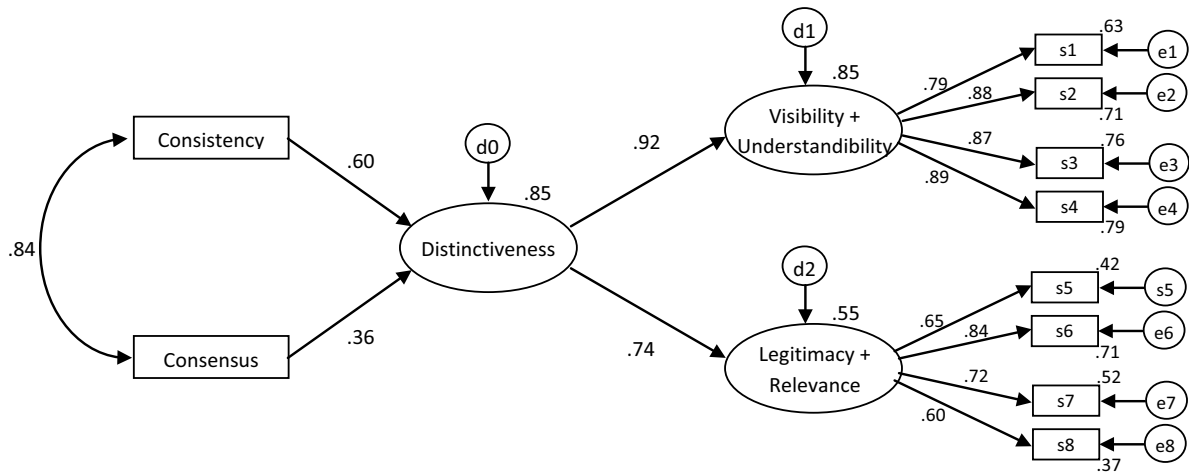


Table 5 shows goodness-of-fit indicators for SEM results. The model titled “Study 2b” in the table shows what happens when ‘legitimacy’ and ‘relevance’ are merged together; “Study 2a” in table 5 represents the model shown in figure 3b.

Table 5 – HR System: SEM Results from studies 1 and 2

Indicators	Study 1	Study 2a	Study 2b
CMIN	38.540	76.602	81.515
DF	34	50	51
CMINDF	1.134	1.532	1.598
CFI	.993	0.966	0.96
RMSEA	.039	0.078	0.083
BCC	86.463	142.440	145.002
BIC	133.036	201.968	202.403

Goodness-of-fit indicators in both studies are within the parameters suggested in the literature (Arbuckle, 2006) to accept SEM models. Figures 3a and 3b have a similar pattern of quality indicators, although figure 3a is slightly better than its counterpart 3b.

From both figures 3a and 3b, a number of important findings need to be highlighted:

- Consistency and consensus reveal a high level of association between one another; also, empirically-generated structures failed to give support to Bowen and Ostroff’s (2004) five attributes of ‘instrumentality’, ‘validity’, ‘consistent HR messages’ (all from consistency), ‘agreement among message sender’, and ‘fairness of the HR system’ (both from consensus);
- Consistency and consensus have a strong impact on distinctiveness (see squared multiple correlations: 0.85 in study 1 and 0.97 in study 2), with a prevalence of consistency (standardised effects of 0.60 in study 1 and 0.78 in study 2, against 0.36 and 0.26 for consensus, respectively for studies 1 and 2);
- The attributes which comprise distinctiveness partly emerged in the data, especially in the model from study 2. In both studies, visibility and understandability are shown together, which suggests that these constructs are in fact very similar and may not even be possible to distinguish.
- Legitimacy and relevance also share a common ground, although data also suggests that these are separate and independent constructs, as shown in figure 3a. Several sentences emerged very clearly in both studies, in support of the variables visibility, understandability, legitimacy, and relevance. Although improvements were introduced

from study 1 to study 2, several sentences still presented problems. Out of the more than 100 sentences tested, only a dozen showed a good support for part of the model. Altogether, they allow the attributes of visibility/instrumentality, legitimacy, and validity, and the dimensions of consensus, consistency, and distinctiveness to be measured.

5. DISCUSSION

5.1 Findings and limitations of the research

This work presented an attempt to operationally define the nine meta-attributes of HR System Strength as proposed by Bowen and Ostroff (2004). A second goal was to explore these authors' model as far as its construct validity is concerned. The number of respondents did not allow deeper relationships to be investigated, but results are interesting enough to stimulate further research. The fact that two studies were conducted in relatively independent conditions, and revealed a similar pattern of results, should offset the effect of the small number of cases used in both cases.

The two studies were designed to account for several measurement issues which have been addressed in the literature on strategic HR, in general, and in the literature on HR Strength, in particular. One such issue is the concept of the HR System itself, which may be represented as the following question: on what grounds and content should people be asked to produce their judgements regarding the HR function? Should they evaluate one or more specific HR practice, and/or should they call to view HR as a whole entity? Study 1 dealt with the first part of the problem, while study 2 aimed to address the second part. Data from the two studies show similar results, which gives support to the idea that HR content and HR process are two distinct concepts, as proposed by Bowen and Ostroff, in 2000 and again in 2004.

This is an interesting finding, since it suggests that the HR function needs to carefully address the *way* it communicates with employees, on top of the *content* of its messages. Sensemaking and causal attribution are complex psychological processes which are affected by both the content of what is perceived and the way it is conveyed to people (Kelley, 1973; Mischel, 1973; Weick, 1995). By its very nature, the HR function has a central communication role in all organisations, hence it should pay attention to this key process and all its features, including communication means, channels, and barriers. Relying on the message as the sole factor shaping individuals' perceptions is just half of the work needed to influence people's interpretation of their surroundings. Future works need to address in full the implications of these findings using Kelley's three-dimensional model. For example, Hewstone and Jaspars (1988) found that in conditions of high consensus, distinctiveness is likely to have a stronger impact on people's readings of the environment. Will this finding hold when the influencing factor is the HR System?

Reliability results of the new scales are encouraging, suggesting that the procedure which was followed to generate and check items was appropriate. Exploratory and confirmatory factor analysis showed a mixed, yet favourable set of results supporting some of the model's predictions.

In the first place, the three dimensions on which the authors built their framework is supported in both studies. In fact, distinctiveness, consensus and consistency appear to be independent constructs, perfectly captured by the sentences developed in this research. This result also supports Kelley's (1967, 1973) attribution theory, which seems to be useful in explaining how an HR setting (HR practices and HR goals) is part of the environment affecting people's interpretation and sensemaking processes. Second, the relationships between the three constructs also seem to follow a pattern expected in the literature. For

example, Bowen and Ostroff (2004) assert that “consistency and consensus are distinct but interrelated concepts” (p. 212). Both in figures 3a and 3b this premise is strongly supported, with a high correlation level between the two dimensions. Another interesting result that runs in favour of the model is the pattern and sequence of relationships between the three dimensions. As proposed by Bowen and Ostroff (2004), and shown in other literature from attribution theory (e.g. Hewstone & Jaspars, 1988), dimension ‘distinctiveness’ “drives up attention” (Bowen & Ostroff, 2004, p.214). Both studies carried out by Hewstone and Jaspars (1988) clearly show that attribution is especially affected by the way people understand their environment, as well as how they see it as visible and relevant. Consensus and consistency are also relevant, but mainly in affecting distinctiveness. These relationships were confirmed in both studies performed in this research. In fact, distinctiveness seems to be dependent on both consistency and consensus, with a stronger weight for consistency, which is also supported by Hewstone and Jaspars’ (1988) findings.

A disappointing set of results derives from the new attributes put forward by Bowen and Ostroff (2004). Neither exploratory nor confirmatory factor analyses provided definite evidence in support of the nine attributes. There are several explanations for this finding. A first explanation is related to the quality of the current research itself, whereas a second one is concerned with the constructs’ definition offered by Bowen and Ostroff (2004).

In the first group of explanations, sample size and sampling procedures are on the first line. In fact, although two independent studies were used to investigate the phenomenon under discussion, and although extreme care was placed on the initial phases of the questionnaire construction, data collection was limited to six medium-size companies with a relatively small number of respondents (fewer than 200 altogether). To fully validate a new instrument, a larger sample and other psychometric techniques are required (Nunnally & Bernstein, 1996). Nevertheless, the main goal of this research – to deliver an initial set of impressions regarding a new instrument to measure HR System Strength – was fully accomplished. The fact that it was carried out in real settings (organisations with HR Systems in place) is also a positive point worth mentioning.

Another shortcoming of the current research is concerned with the level of analysis. Bowen and Ostroff’s (2004) model is interesting because it offers the possibility of studying phenomena at different levels of analysis. In the present investigation, data was collected only from individual employees. Due to small sample size, it was not possible to collapse individual data into group data. Such kind of data would satisfy some of the recommendations described by James (1982) and Chan (1998) to perform good-quality multi-level research. Central-tendency measures are intrinsically interesting (as recognised by Bowen and Ostroff themselves), but some sort of dispersion measures would also allow extra information on some of the attributes and dimensions, such as consensus and consistency.

Finally, although perceptions of people are key to evaluating how strong a HR System is, it is possible that for some attributes, people do not have enough information to produce meaningful answers. An obvious example is recruitment and selection: after working for more than 2 or 3 years in the organisation, regular employees (i.e. non-HR professionals) would probably have forgotten many of the details of the process in which they enrolled when they entered their company. In such a case, they would not be able to deliver a valid answer to many of the questions on the questionnaire. One can argue that performance appraisal is more present in an employee’s working life, but in reality, no one really knows whether or not an employee has enough information to produce a meaningful and valid answer regarding a particular HR practice or set of practices.

One way to minimise this problem is to use multiple sources of information (triangulation of information sources; Jick, 1979). In HR System Strength research, this may require information gathered both from employees and their immediate supervisors. Some of

the nine attributes may even be easier to collect from immediate supervisors than from employees.

With regard to the second group of explanations, some of the attributes put forward by Bowen and Ostroff (2004) may require further definition and delimitation. This unsatisfactory or inadequate definition may in some cases have caused the problems detected with the data. For example, the attribute 'consistent HR messages' is defined at three levels (Bowen & Ostroff, 2004, p. 211): a) between senior managers' explicitly stated goals and values, and what employees actually conclude those goals and values to be based on their perceptions of HR practices; b) internal consistency among the HR practices themselves; and c) stability over time. On the other hand, the attribute 'validity' (also from the Consistency dimension) is defined as (p. 211) consistency between what HR practices claim to do, and what they actually do. This definition of 'validity' is very similar to the first interpretation (a) of 'consistent HR messages'. One might hypothesise that for employees at low levels in large organisations, it may be difficult to distinguish between different message senders. In other words, for some employees it is probably a minor issue whether messages regarding appropriate organisational behaviours come from senior managers or from HR practices.

The relationships between the nine attributes also need clarification. Although this stream of research on HR is very much in its initial stages, it is necessary to clarify how the attributes relate to each other. Several hints are advanced by Bowen and Ostroff (2004). From an intuitive standpoint, connections between attributes are also expected. For example, understandability of a particular HR practice (knowing how it works, its criteria, and so on) may be highly dependent on its visibility or salience; in fact, how can one generate understanding of something which is not visible? This may also help explain why understandability and visibility merged together in the data in both studies. On the other hand, the practical challenges associated with specifying several connections between the nine attributes, hence producing a non-orthogonal model, are of several types and difficulties. For instance, an extremely large sample would be required to test with some quality a model with three dimensions, nine attributes, and a dozen or two paths amongst them.

5.2 Limitations and future research

The exploratory nature of the current work raises more questions than answers, but it also stimulates thinking both at the theoretical and methodological levels. A few questions which need to be addressed in the future are:

- Should aggregate data be collected at an organisational level or group level? Culture theorists (e.g. Schein, 1981) accept that sub-cultures within organisations may exist, suggesting that smaller groups than the organisation can be used to carry out studies; likewise, Bowen and Ostroff (2004) suggest that people working in small groups may share, to a greater or lesser extent, a particular perspective regarding HR practices.
- Is HR Strength the sole influencing factor in shaping people's view of the situations? Yukl's (1998, p.5) definition of leadership suggests that HR Strength may not be alone in the influencing process: leadership is "a process through which an organisation's individual or group is able to influence all other members' interpretation of events". Future work on the influencing factors shaping one's perspective of the situation needs to take leadership into account.
- Last but not least, the whole idea of the HR Strength and Situation Strength concepts is to propose a way to link the HR function with individual and organisational performance (Bowen and Ostroff, 2004). As far as the current research was concerned, such a link was not a goal; therefore no final conclusions can be drawn regarding the validity of the model. Future work needs to test this theory, by incorporating individual and organisational performance variables and measures.

REFERENCES

- Aijala, A., Walsh, B. and Schwartz, J. (2007). Aligned at the Top: How business and HR executives view today's most significant people challenges – and what they're doing about it. Special report of *The Economist Intelligent Unit and Deloitte Touche Tohmatsu*, Deloitte Development LLC.
- Arbuckle, J.L. (2006). *Amos 16.0 User's Guide*. Spring House, PA: Amos Development Corporation.
- Bowen, D.E. and Ostroff, C. (2004). Understanding HRM-Firm performance linkages: The role of the "strength" of the HRM system. *Academy of Management Review*, 29(2), 203-221.
- Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, 83(2), 234-246.
- Chen, S.-J., Lin, P.-F., Lu, C.-M., and Tsao, C.-W. (2007). The moderation effect of HR strength on the relationship between employee commitment and job performance. *Social Behavior and Personality*, 35(8), 1121-1138.
- Combs, Y.L.J., Hall, A. and Ketchen, D. (2006). How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology*, 59, 501-528.
- Delery, J.E. and Doty, D.H. (1996). Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *Academy of Management Journal*, 39(4), 802-835.
- Dorenbosch, L., de Reuver, R. and Sanders, K. (2006). Getting the HR message across: The linkage between line-HR consensus and "commitment strength" among Hospital Employees. *Management Review*, 17(3), 274-291.
- Edgar, F. and Geare, A.J. (2005). Employee voice on human resource management. *Asia Pacific Journal of Human Resources*, 43(3), 361-380.
- Ferris, G.R., Arthur, M.M., Berkson, H.M., Kaplan, D.M., Harrell-Cook, G. and Frink, D.D. (1998). Toward a social context theory of the human resource management-organization effectiveness relationship. *Human Resource Management Review*, 8(3), 235-264.
- Gerhart, B., Wright, P.M. and McMahan, G.C. (2000). Measurement error in research on the human resources and firm performance relationship: Further evidence and analysis. *Personnel Psychology*, 53(4), 855-872.
- Hewstone, M. and Jaspar, J. (1988). Implicit and explicit consensus as determinants of causal attribution. Two experimental investigations. *European Journal of Psychology*, 18(1), 93-98.
- Ichniowski, C., Shaw, K. and Prennushi, G. (1997). The effects of human resource management practices on productivity: A study of steel finishing lines. *The American Economic Review*, 87, 297-313.
- James, L.R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology*, 67, 219-229.
- James, L.R., Demaree, R.G. and Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69(1), 85-98.
- Jick, T.D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24, 602-611.
- Kelley, H.H. (1967). Attribution theory in social psychology. *Nebraska Symposium on Motivation*, 15, 192-238.

- Kelley, H.H. (1973). The processes of causal attribution. *American Psychologist*, 28, 107–128.
- Mischel, W. (1973). Toward a cognitive social learning conceptualization of personality. *Psychological Review*, 80, 252-283.
- Nunnally, J.C. and Bernstein, I.H. (1996). *Psychometric Theory* (3rd ed.). McGraw-Hill.
- Ostroff, C. and Bowen, D.E. (2000). Moving HR to a higher level: HR practices and organizational effectiveness. In K.J. Klein and S.W.J. Kozlowski (Eds.), *Multilevel theory, research and methods in organizations: Foundations, extensions, and new directions* (pp. 211-266). San Francisco, CA: Jossey-Bass.
- Payne, R.L. (2000). Culture and climate: How close can they get? In N.M. Ashkanasy, C.P.M. Wilderom and M.F. Peterson (Eds.), *Handbook of organizational culture and climate*, 163-176. Thousand Oaks, CA: Sage.
- Pfeffer, J. (1994). *Competitive advantage through people: Unleashing the power of the work force*. Boston: Harvard Business School Press.
- Podsakoff, P.M. and Organ, D.W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12, 69-82.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. and Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903.
- Rego, A. (2000). Justiça organizacional – Desenvolvimento e validação de um instrumento de medida. *Psicologia*, XIV (2), 285-307.
- Rego, A. (2002). Na senda da compreensão das percepções de justiça dos professores do ensino superior. *Psychologica*, 29, 39-51.
- Schein, E.H. (1981). Coming to a new awareness of organizational culture. *Sloan Management Review*, Winter, 3-16.
- Schneider, B., Salvaggio, A.N. and Subirats, M. (2002). Climate strength: A new direction for climate research. *Journal of Applied Psychology*, 87, 220-229.
- Ulrich, D. and Brockbank, W. (2005). *The HR Value Proposition*, Boston, MA: Harvard Business School Press.
- Ulrich, D. (1997). *Human resource champions: The next agenda for adding values and delivery results*. Boston: Harvard Business School Press.
- Weick, K.E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.
- Yukl, G. (1998). *Leadership in Organizations* (4th Ed.), Englewood-Cliffs, NJ: Prentice-Hall.