

Best Companies Rankings Evaluation:

A serious branding exercise or a
spectacular nonsense?





Best Companies Rankings evaluation: A serious exercise in branding or a spectacular nonsense?

A preliminary assessment

N. J. Higgins & G. Cohen 2007

This paper provides a preliminary evaluation of the Best Companies Rankings with specific focus on the survey construct and index methodology, evaluating evidence based on available published data.

International School of Human Capital Management

*2nd Floor
Berkeley Square House
Berkeley Square
London
W1J 6BD*

*Tel: +44 (0)207 887 6121
Fax: +44 (0)207 887 6100*

*www.ISHCM.com
e-mail: info@ISHCM.com*

Content

Best Companies - introduction	8
Is Best Companies about employee engagement? A review of the published methodology	9
(a) Survey response biases and the reliability of question statements (Types I –XIV)	12
(b) Best Companies BC16 core question-statement set reliability (construct) analysis	14
(c) Best Companies expanded question-statement set reliability (construct) analysis summary	15
(d) BC16 and factor reliability	16
Comment on the underlying factor model	17
Factor analysis and its limitations	18
Garbage in, Garbage out: The fallacy of the 'Red hat-Red car' experiment	19
A review of the core process	20
Gaining Best Companies accreditation	21
The Sunday Times-Best Companies Rankings	22
Is the monicker 'best company' an accurate reference term?	22
Benefits to organisations participating in the Rankings	23
Issues for organisations participating in the Rankings	23
Best Companies - funding	24
Final comment	24
Bibliography	25
Appendix I	26

Executive summary

Concerns over Best Companies construct and methodology

- 1 The evidence points to the fact that Best Companies is a clever marketing scheme which has appeal to organisations, particularly SMEs looking to differentiate their employer brand.
- 2 However, the core survey construct used, in terms of the questions, is fundamentally flawed. This effectively raises serious doubts as to the reliability and ultimate validity of the data produced for the Sunday Times BC Index, the Best Companies accreditation and associated employee engagement construct.
- 3 This is further compounded in that Best Companies actually ignores its own advice on best practice design as set out in its Methodology document as published.
- 4 Only four of its core set of 16 questions (BC16) pass the reliability test and over half of the remaining 50 plus questions in the expanded version suffer from the same fate.
- 5 Thus, despite any 'valid' research attempts to utilise the data (such as factor analysis) there remains the large probability of suffering from '**Garbage in-Garbage out**' (GIGO) syndrome, rendering any potential insight as spurious.
- 6 The fact that Best Companies state that they did not provide an original underpinning theoretical or experiential framework to their work, preferring to 'discover it' is at odds with normal multivariate/factor analysis protocol.
- 7 It is also worthy of note that Best Companies methodology has only recently been published (June 2006) whereas the initial exercise was introduced in 2001.
- 8 There are also concerns over the actual anonymity of the overall process as it is not strictly 'anonymous'.
- 9 With regard to organisation employee populations, there are concerns over the use of sampling data and its limitations versus empirical data for the whole respondent employee population with regard to any ranking exercise and particularly any subsequent organisational intervention.

- 10 Given the evidence of the flawed question design and the ignoring of their own best practice advice, coupled with the bizarre research explanation, one can only conclude that the BC Methodology has been retrospectively 'fitted' and published. This raises serious issues of credibility.

Credibility of The Sunday Times-Best Companies Rankings

- With regard to the various Rankings, we observe that the annual 'churn' of organisations appearing in the rankings is around 50% meaning that there is effectively a new index every two years.
- Coupled with the fact that published organisations represent less than 0.7% of organisations in the UK it is hard to provide any credence or meaning to the rankings.
- The endorsement of The Sunday Times and DTI to the rankings (and by association any related interaction with organisations) provides a false air of authenticity, raising questions around such endorsements.

Other general comments

- We acknowledge that the ST-BC Rankings may have raised the profile of certain people management issues particularly for SMEs (notwithstanding the existing contribution of Investors in People) but would question as to the drivers for Rankings involvement and whether they are truly people-centric.
- It should be noted here that the Scheme is not open to organisations with less than 50 employees and less than 5 years trading thus bringing into question the actual SME focus.
- There is very little evidence available as to the benefits gained by those organisations appearing in the Rankings, save for their public exposure. Given the 'recruitment branding purpose' with which the Rankings were originally published, we would question as to the local benefits any small companies receive (who make up a large proportion of entrants).

Final note of caution

- For entrants or users of the Best Companies scheme we would therefore urge caution in the form of 'caveat emptor' – 'let the buyer beware'.

Best Companies - introduction

The Best Companies Rankings, were introduced to replicate the US 'American Best Companies to Work for' list, and have been published annually in the Sunday Times since 2001.

The following table shows how the scale and focus of the Best Companies Rankings has evolved from this initial publication.

	Total	Small	Mid	Large
2001	50	50		
2002	100	100		
2003	100	100		
2004	150	50	100	
2005	210	100	100	10
2006	220	100	100	20
2007	220	100	100	20

'Special Awards' were introduced in 2005 in acknowledgement of organisations scoring highly in particular areas of assessment. These include the eight 'factors' underlying the survey methodology, in addition to awards that have been sponsored by the DTI, Learning and Skills Council, Investors in People, and various private companies, charities and institutes/foundations.

The Sunday Times supplements, that include the publication of the rankings, feature short profiles of the ranked companies in addition to specific features. These features are typically authored or perspectives put forward by employees of Best Companies, the DTI and the Sunday Times.

The company ranking profiles include selected quantitative demographic data¹ with a short narrative (and frequently a photograph). The narrative primarily focuses on staff benefits, training and community involvement, areas of high response in the questionnaire and comments/interviews with existing employees.

We note, however, that the quantitative and narrative aspects can throw up apparent contradictions. For example, certain organisations featuring in the Best Companies rankings report annual staff turnover rates in excess of 30%, calling into question the extent to which the organisation is in reality perceived as a 'Best Company' and the true value of providing particular staff benefits.

¹ Annual sales, Staff numbers, Male:Female ratio, Under-35s/Over 55s, Staff turnover, Proportion of workforce earning over £35,000 and Typical job.

Is Best Companies about employee engagement? A review of the published methodology

Introduction

Best Companies publish a definition of employee engagement on their website as follows:

Engagement can be defined as -

“An employees (sic) drive to use all their ingenuity and resources for the benefit of the company”.

Best Companies has one publication available from its website, entitled ‘Best Companies methodology’.

This was published in 2006, five years after the release of their inaugural index ranking. The document explains their survey methodology ‘in detail’² both for the creation of the Sunday Times lists and the Best Companies accreditation system.

The document provides an overview of five key components:

1. **Surveying employees** - which outlines the best practice design of an employee questionnaire and the key issues/things to avoid as well as the Best Companies process
2. **Developing a theoretical model of workplace engagement** – which outlines the use of ‘factor analysis’ of data and the Best Companies factor model
3. **Measuring the performance of organisations** – which describes the Best Companies process of assessment
4. **Scoring the survey** - which provides further context around factor weightings
5. **Accreditation methodology** – which outlines the process around the Best Companies accreditation

Interestingly and confusingly, the introduction states that:

“Over the last six years Best Companies has developed and tested what we believe is the most accurate and valid survey instrument in the UK for measuring employees attitudes to their work and their organisation.”

Thus, there is a discrepancy in the use of the terms employee attitudes and employee engagement which are not the same, and there is no mention of the employee engagement definition as published on the website.

Questions around approach

Ordinarily, with methodologies of this sort in scientific research, we would expect to critique the research hypothesis and underpinning model, prior to the evaluation of any instruments or survey constructs used in the collation of data before evaluating the findings and further outputs (see the appendix for a general schematic). Most commonly there are nine steps.

Unusually, and somewhat questionably Best Companies appear to have started their underlying methodology half-way through (i.e. step 5) as there is no actual methodological explanation of the underlying ‘engagement framework’, relying instead on reference to the initial question-set.

Best Companies provide an explanation for this:

“The factors in the model are totally derived from the employee data with no personal subjective input. Thus, the factors in the model... are derived purely statistically from those things which employees score as most important to their engagement. These engagement factors were not “thought up” by Best Companies staff, or any external experts, they are a research finding from the data.”

Best Companies Methodology document, June 2006, Section 2.

Thus, Best Companies appear to retrospectively say that they did not start with an underlying model but establish one from the derived data (through factor analysis).

² Unfortunately there are no page numbers contained in the document to reference.

This therefore leads to the question that if this was the case, how did the first set of questions evolve without some hypothesis?

We can conclude that either Best Companies statement is untrue, and there was an underpinning hypothesis reflected in their initial survey construct or that they have committed a scientific faux-pas by not carrying out the previous four fundamental steps to scientific research.

This is particularly key in qualitative approaches such as that represented in themes such as employee engagement.

Understanding the methodologies in play

However, before looking at the question design, there is a need to understand that there are three distinct research methodologies at play here:

- i. The methodology and rationale used to create the underlying employee engagement model from previous or proprietary research
- ii. The methodology underpinning the question-statement³ design which is used to collate and then populate the engagement model, and which needs to be valid to avoid problems associated with 'Garbage in, Garbage out syndrome'.
- iii. Various multivariate methods and statistical techniques to analyse and validate the data

As we have already noted, Best Companies did not undertake the first core methodology preferring to use factor analysis⁴ to, what neutral observers may consider dubiously, 'discover' it. Thus the Methodology document focuses on (ii) and (iii) above.

The obvious danger as pointed out in any standard texts on qualitative research (particularly psychology-related) is the missing of defined model dimensions and/or key inputs/variables thus rendering any collation

exercise subject to skewing or resulting in meaningless data.

Scoring scale in use

Best Companies advocates the use of a 7-point Likert⁵ scale which, as we would advocate, is a good scale to use (on the proviso of good questions sets). The scale in use is as follows:

- Strongly agree
- Agree
- Slightly agree
- Neither agree nor disagree
- Slightly disagree
- Disagree
- Strongly disagree.

However, we would reiterate that the Likert scale will only be as good as the individual question design allows.

Question design

The Best Companies methodology document makes use of previous research in question-statement design, quoting good reference texts.

It states '...valid question development is crucial in producing valid and reliable data, and because we are often asked why it is not possible to just add other questions to our survey.....Developing good survey questions depends first on having a good background knowledge of the area and issues you want to measure.....the wording of questions is crucial and changing a single point of either content or context may introduce a bias into the answers.'

The methodology text quite rightly points out the potential design flaws and faults with regard to question design. Survey response biases are created through flawed design and will seriously impair the validity of any subsequent data collation and inferential analysis such as factor analysis.

The document clearly states that *'Best Companies employee survey questions are designed as far as possible to avoid these potential biases.'*

This is to all extent common sense, particularly given the reliance that Best Companies places on factor analysis at the expense of any

³ We use the term 'question-statement' as many survey questions are in fact statements in order that they can generate varying responses as per the response scale used

⁴ There is a sizeable piece of literature devoted to factor analysis in terms of its application and limitations

⁵ Standard technique for garnering responses to qualitative questions. Many use 5-point scale i.e. two positive, two negative and one neutral response generating. 7-point increases this to three positive and negative choices.



underlying framework. The reliability of the question statements becomes paramount.

A typology of question design flaws

As useful knowledge to all practitioners we have repeated these here with some amendments, adding our own typology where it provides additional value. They are Type-:

- I. Leading questions
- II. Double barrelled/multiple questions
- III. Knowledge or projection (proxy)
- IV. Response extremity
- V. Responses open to social desirability/prestige
- VI. Responses implying causality
- VII. Questions that impose unwarranted assumptions
- VIII. Questions that include hidden contingencies
- IX. Questions that include ambiguous time periods
- X. Questions containing concepts that are open to differing interpretation
- XI. Question that duplicates another or is a reverse of another
- XII. Questions requiring a tendency to acquiesce and/or imply 'psychological threat or hostility
- XIII. Questions that are exclusively positively or exclusively negatively clustered
- XIV. Questions which are culturally loaded and or overly long

This list is not exhaustive, but contains the main issues when designing reliable questions that provide good valid data.

Best Companies refers to its core set of 16 questions⁶ which provides the basis for assessment for both index rankings and accreditation.

It is therefore unfortunate to find that these core sixteen questions contain many type flaws that are identified above. Best Companies appear to have ignored their own (and others) best practice advice.

Given the clear irrevocable statement made by Best Companies above, this has serious

repercussions for the entire validity⁷ of the 'Best Companies' construct, Rankings, Accreditation and any subsequent 'organisational' intervention.

The Methodology document does state that a pilot process began in 2001 with 134 questions which were put through an extensive pilot process and 'statistically tested' (though no details are made available). The commentary curiously stops at this point, making a pointed remark regarding hindsight and response bias but no further information is forthcoming.

If this is true this surely raises even further questions as to the survey construct given the information provided on the next pages.

The comment that 'the statistical testing...on each year's new data....has effectively given us five totally independent statistical analyses of the survey materials' contains two notable issues:

- The statement is erroneous in that the 'statistical tests' are in fact 'repeat' tests which are not 'totally independent'
- The phrase 'statistical analyses' does not mean 'tests' as the message of the paragraph implies.

The following pages provide a brief evaluation of the questions used in the Best Company exercise:

- a. Survey response biases and reliability (Type I-XIV)
- b. BC16 core set reliability analysis
- c. BC expanded set summary reliability analysis
- d. BC16 and factor reliability

⁶ There are a further 50 plus statements whose responses are used for factor analysis and subsequent interpretation/actioning

⁷ Our definition here is the 'layman' term of valid i.e. opposite to invalid; as opposed to the various technical definitions that are in existence in the area of research

(a) Survey response biases and the reliability of question statements (Types I –XIV)

	Q-S RELIABILITY: TYPES TO AVOID	EXPLANATION AND EXAMPLES
i	Leading (loaded) questions	Questions that 'lead' responses and thus bias; for example, 'I am under too much pressure to perform well' implies that pressure adversely affects performance or 'Profit is the only thing driving this organisation' which would be better suited to a multiple choice type question that contains all other 'things' for selection.
ii	Double barrelled/multiple questions (erroneously referred to as 'single question' in the BC document)	A question which actually contains two questions thus rendering responses invalid; for example, 'Senior managers do a lot of telling but not much listening' . Taking a response such as 'I agree' – does that mean the respondent agrees with both 'do a lot of telling' and 'not much listening'; or 'do a lot of telling' only, or 'not much listening' only? Similarly for disagree. To compound the type error it is also a doubly-loaded assumptive (Type II) question-statement.
iii	Knowledge or projection (proxy),	i.e. Questions that project as a proxy for others (conjecture) or asks about how somebody else feels or indeed asks about feelings rather than asks definitively providing further vagueness of subjectivity of response; For example, 'Profit is the only thing driving this organisation' (notwithstanding the underlying assumption of 'profit focus is bad'). Also consider 'I am proud to work for this organisation' instead of 'I feel proud to work for this organisation' .
iv	Response extremity	Questions which limit the response range and/or invalidate response interpretation. For example, 'I am under too much pressure at work to perform well' or 'I have a great deal of faith in the person leading this organisation' essentially limit responses to yes/no. It is very difficult to interpret or differentiate responses such as the opposites 'I slightly agree' and 'I slightly disagree' with the main statements (notwithstanding any other type issues).
v	Responses open to social desirability and/or prestige	Questions which focus on an individual, and or status/cause, for example, 'My organisation makes a positive difference to the world we live in' (notwithstanding the definitional issues contained therein)
vi	Responses implying causality	These questions should be avoided for their underlying bias (this is in fact what post-survey analysis is used for) to response interpretation 'My health is suffering because of my work' .
vii	Questions that impose unwarranted assumptions	These can imply bias in responses. For example, 'I am under too much pressure at work to perform well' which assumes that 'pressure' inhibits performance (notwithstanding the interpretation of pressure). Also questions such as 'this organisation is run on strong values/principles' which already assumes the respondent identifies that these are the right ones (to become a meaningful question).
viii	Questions that include hidden contingencies	These are where questions can only be answered by a particular subset of available population, i.e. questions that may refer to, for example, profit and then being asked of public or NFP sector employees such as 'Profit is the only thing driving this organisation' .



Survey response biases and the reliability of question statements (continued...)

	Q-S RELIABILITY: TYPES TO AVOID	EXPLANATION AND EXAMPLES
ix	<i>Questions that include ambiguous time periods</i>	i.e. Questions that assume uniformity across time in response which may be misrepresentative without further definition, for example, 'My work is stimulating' (some of the time, all of the time, last week, last three months, or previous 12 months?). This can quite strongly effect the response given without a timeframe.
x	<i>Questions containing concepts that are open to differing interpretation</i>	These questions effectively confuse meaning and thus cause problems with response interpretation and reliability i.e. questions containing buzzwords or phrases which have different/broad interpretations to respondents; for example, 'Working in this team gives me a buzz' , or 'My organisation makes a positive difference to the world we live in' (notwithstanding other type issues).
xi	<i>Question that duplicates another or is a reverse of another</i>	The question itself may not be the issue but the duplication (or reverse duplication) may well be, owing to the issue of double-counting or -representation. A 'reverse-pairing', for example, are 'My work is stimulating' and 'I am bored with the work I do' (other type-problems notwithstanding). An example of a 'congruent pairing' would be 'I love working for this organisation' and 'I am proud to work for this organisation' (again notwithstanding any other related reliability-type issues). Reverse/congruent pairings should not be contained within a 'core' question set due to the potential for manipulation or skewing of results.
xii	<i>Questions requiring 'tendency to acquiesce' and/or imply 'psychological threat or hostility'</i>	(Normally in-house surveys are most at risk). However, in small samples (and particularly small organisations) the question 'I have a great deal of faith in the person leading this organisation' implies a tendency for respondents to acquiesce. (For large organisations, this question infers less 'tendency to acquiesce' but its individual slant becomes much less meaningful).
xiii	<i>Questions that are exclusively positively or exclusively negatively clustered</i>	This issue is more to do with the design of the overall question set (referred to in the text as 'Yea/Neh' saying and 'response contraction') which invoke potential response bias if care is not taken on question structure.
xiv	<i>Questions which are subject to issues of culturally loaded and or overly long</i>	(Similar to 'differing interpretation' in (x) above but more contextual rather than definitional)

Evaluation of question wording

The following page reviews the BC16 core questions to ascertain the degree to which Best Companies have applied their methodology to the core questions identified. For the purposes of this critique, we have reviewed the core BC16 set as defined. The findings are representative of our review of the remaining 50+ questions (forming the expanded set).

(b) Best Companies BC16 core question-statement set reliability (construct) analysis

	BC16 Question-statement	Reliable construct	TRANSGRESSION
1	I have a great deal of faith in the person leading this organisation	✗	Fails on THREE counts
2	Senior managers of this organisation do a lot of telling but not much listening	✗	Fails on FOUR counts
3	My manager cares about how satisfied I am in my job	✓	(if ignore for any social desirability/ acquiescence bias)
4	My manager helps me fulfill my potential	✓	(if ignore for any social desirability/ acquiescence bias)
5	My work is stimulating	✗	Fails on TWO counts
6	I am bored with the work I do	✗	Fails on TWO counts
7	I am under too much pressure at work to perform well	✗	Fails on FOUR counts
8	My health is suffering because of my work	✗	Fails on THREE counts
9	My team is fun to work with	✗	Fails on THREE counts
10	Working in this team gives me a buzz	✗	Fails on THREE counts
11	Profit is the only thing driving this organisation	✗	Fails on FOUR counts
12	My organisation makes a positive difference to the world we live in	✗	Fails on THREE counts
13	I love working for this organisation	✗	Fails on TWO counts
14	I feel proud to work for this organisation	✓	Reliable (more so if 'am' replaces 'feel')
15	This organisation is run on strong values/principles	✗	Fails on TWO counts
16	I feel I receive fair pay for the responsibilities I have in my job	✓	Reliable (again more so if 'I feel' is dropped from the sentence)

Summary comment

- Thus of the sixteen core questions underpinning the employee engagement construct, index ranking and accreditation, only four pass the reliability test (allowing for adjustments).
- Any question failing two or more type issues should be discounted completely from any subsequent analysis or assessment due to the unsafe reliability.
- Best Companies do not publish any related statistical reliability test data (e.g Cronbach alpha, Chi-Square, 'goodness-of-fit index' or AGFI as guides).
- The real danger is that given the flaws in question-statement design, any subsequent analysis and employee engagement is very exposed to 'Garbage in, Garbage out syndrome'.
- This is particularly acute if any subsequent organisational intervention is driven on the back of this data.
- The design flaws are perhaps even more startling given that Best Companies seemingly ignore their own advice contained within the Methodology document as published.

(c) Best Companies expanded question-statement set reliability (construct) analysis summary

Factor relating to Question-statements	Reliability of FACTOR	Question-statement (cluster) reliability
Leadership	✘	SIX (out of ten) fail reliability test
My manager	✘	FIVE (out of ten) fail reliability test
Personal growth	✘	TWO (out of five) fail reliability test
Well being	✘	SIX (out of ten) fail reliability test
My team	✘	FIVE (out of eight) fail reliability test
Giving something back	✘	FOUR (out of four) fail reliability test
Company	✘	TWO (out of six) fail reliability test
Fair deal	✘	SIX (out of twelve) fail reliability test

Summary comment

- Of the 65 questions reviewed, over half were judged to be unreliable. This allowed for instances where better wording would have increased reliability.
- A number of questions have what seems to be a very tenuous link with the concept of employee engagement at this level.
- Again, Best Companies do not publish any related statistical reliability test data (e.g. Cronbach alpha or Chi-square as guides) for the 'factor cluster-sets'.
- Given the degree to which design flaws populate the overall instrument, subsequent analysis around employee engagement and planned organisational intervention would be extremely suspect.
- Again, reiterating previous, the design flaws have ignored the advice of Best Companies' own Methodology document as published.

(d) BC16 and factor reliability

	Question-statement	Reliability construct	'Engagement' Factor	Reliability of 2Q construct	Reason
1	I have a great deal of faith in the person leading this organisation	✗	Leadership	✗	Unreliable core questions
2	Senior managers of this organisation do a lot of telling but not much listening	✗			
3	My manager cares about how satisfied I am in my job	✓	My manager	✓	(if ignoring social desirability/ acquiescence)
4	My manager helps me fulfill my potential	✓			
5	My work is stimulating	✗	Personal growth	✗	These are a 'reverse-pairing' effectively meaning either double counting or false correlation
6	I am bored with the work I do	✗			
7	I am under too much pressure at work to perform well	✗	Wellbeing	✗	Unreliable core questions
8	My health is suffering because of my work	✗			
9	My team is fun to work with	✗	My team	✗	Unreliable core questions
10	Working in this team gives me a buzz	✗			
11	Profit is the only thing driving this organisation	✗	Give something back	✗	Unreliable core questions
12	My organisation makes a positive difference to the world we live in	✗			
13	I love working for this organisation	✗	Company	✗	Unreliable core question and potential duplication/double-counting
14	I feel proud to work for this organisation	✓			
15	This organisation is run on strong values/principles	✗	Fair deal	✗	Unreliable core question
16	I feel I receive fair pay for the responsibilities I have in my job	✓			

Summary comment

- Essentially only one factor (My manager) would seemingly provide sufficient reliability of question construct design.
- Again Best Companies do not publish any related statistical reliability test data (e.g Cronbach alpha or Chi-square as guides) for the factors themselves.
- Given the fact that no base model was defined and with it, sufficient definition of each factor, it is difficult to assess the validity of the factors themselves from a construct design perspective.
- The design flaws populating the overall instrument provide sufficient doubt to the robustness of the overall 'engagement construct'.

Comment on the underlying factor model

Unfortunately as Best Companies did not start with an underlying 'engagement' model we are unable to comment accordingly. We are somewhat surprised at this stance given the plethora of connected research available stretching back over some 60 years.

For example in our own pursuit of understanding and defining employee engagement our 'meta-review' included a trawl through over 1500 relevant journal papers and publications covering topics such as:

- High performance work environments
- Work psychology
- Organisation behaviour
- Organisation citizenship behaviour
- Individual work commitment and motivation
- Management & leadership
- Expectancy and valency theories
- Self efficacy
- Organisation design
- Organisation and team performance
- Labour economics.

This provided a rich background to understand the components of employee engagement and to derive a working model (complete with factor analysis and structural equation modelling).

Given the reliability issues surrounding the question set, we will limit our commentary on Best Companies 'emergent' eight factors of engagement to the issue of selection and definition.

In factor analysis, the issue primarily revolves around the naming and categorising of the factors. Though Best Companies factors of engagement may have resonance or be identified by readers superficially, the lack of initial definition and underpinning model means that their subsequent factor analysis leaves them at the mercy of the question design and subsequent data collation.

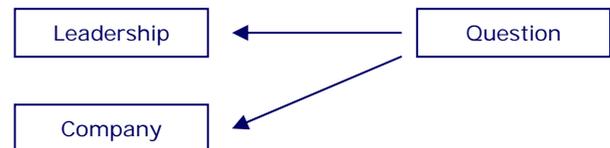
To provide a simple illustration of the challenges besetting the categorisation of factors, we reviewed the question set (allowing for the all of the various Type errors on a theme basis).

One of the drawbacks with the Best Companies method of factor analysis is that we understand each question is assigned to each 'artificial' factor for assessment and scoring, on what is

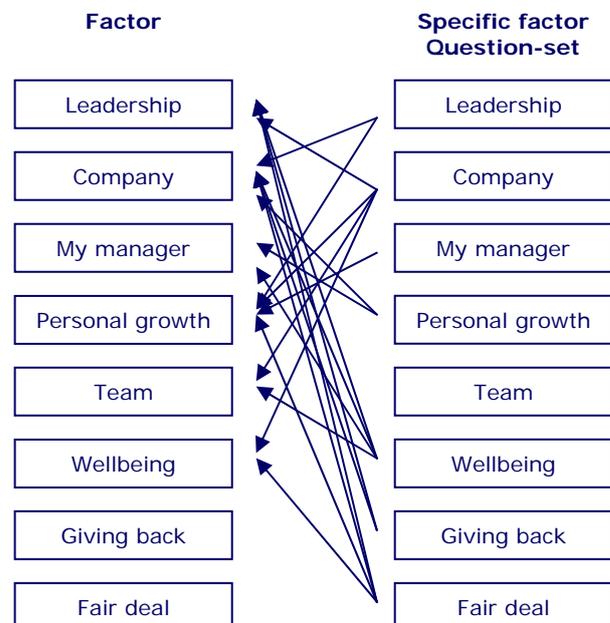
termed a 'unidimensional' basis, i.e. the item question score relates solely to that factor.



The problem however is that many of the questions can be easily interpreted to score under another factor or factors, what is termed multidimensional. Under our analysis, we were able to ascribe 38 questions (c. 60% of the total) that could represent another factor (with 6 being multifactor).



Further the entire question set of two particular factors could be represented elsewhere, meaning potential redundancy of these factors.



—————> Potential factor alternative for an individual question

This finding suggests that there is more than one factor model to represent Best Companies engagement but the question set unreliability means that further analysis is irrelevant.



Factor analysis and its limitations

Best Companies use a particular technique called factor analysis which provides statistical insight and thus statistical validity though appearing not to have tested for question reliability which is extremely odd in this type of qualitative exercise.

For example no Cronbach alpha⁸ statistics are published for either the BC16 or for the questions relating to its eight factors (two question clusters) as identified which provide a guide on the overall reliability of the question set(s) used.

Though factor analysis is a known statistical method it does have limitations and is subject to the 'Garbage-in, Garbage-out' syndrome since factor analysis will always provide correlational factors.

The output is still limited to the initial design. Without this process, there is the danger of circularity in proof. It is therefore deemed unsafe without certain safeguards given that existing parameters and factor analysis correlations can always be found in any ensuing analysis.

Much is made in the methodology document of the use of exploratory and confirmatory factor analysis to isolate the factors of workplace engagement with their underlying questions – indeed this is the statistical technique that the Best Companies methodology effectively stakes its credibility upon. This is because of the admission that the researchers deliberately sought not to impose any pre-conceived framework as a means of exploring the underlying factors of engagement.

Essentially, factor analysis has the purpose of identifying the correlations through a large number of variables by defining the common dimensions underlying it ("factors"). This approach allows a researcher to identify the different variables.

Statistical commentators, however, call into question the application of factor analysis as a means of exploring and confirming the same

factor model. One of the core problems is that factor analysis will always produce factors to varying degrees (see overleaf).

As we have already stated, factor analysis can unfortunately suffer from 'Garbage in – Garbage out' syndrome' and the problem lies with the underlying construct or model. If factor analysis is done in an indiscriminate fashion, i.e. data on a large number of variables are collated and then factor analysis is used to 'figure it out' then the possibility of poor or misleading results is high.⁹

The use of factor analysis does not preclude the requirement of a conceptual basis or underpinning model to effectively structure the data requirement and particularly in qualitative analysis where defined question-statements are required.

"The Best Companies factor model of workplace engagement was initially created with exploratory factor analysis and then checked and refined with confirmatory factor analysis. Each year as new data arrives both these analyses are repeated and new questions checked."

Methodology document, June 2006, Section 2.1.2

⁸ Standard reliability test particularly used in qualitative analysis. It is a statistical measure of response reliability. Like all tests it can be manipulated but remains the most significant in terms of robustness.

⁹ See Hair, Anderson, Tatham & Black, *Multivariate data analysis, Chapter 3 Factor Analysis*, International 5th edition, Prentice Hall



Garbage in, Garbage out: The fallacy of the 'Red hat-Red car' experiment

Suppose that I want to find out something about colours, cars and days of the week. I don't start with any hypothesis but conduct an experiment to find out a connection between wearing a particular coloured hat on a particular day of the week and any other variables that may ensue.

I pick six hats representing the spectrum: red, blue, green, yellow, black and white. I stand for seven days ticking off the various cars that go past with their associated colours. I repeat the process wearing a different coloured hat for each day of the week several times. I collect my data and conduct factor analysis.

Amazingly my results show correlations between several items. For example, the data tells me that on Thursdays, if wearing a yellow hat, I am more likely to find that red cars will be the most common colour. I also find some other data which suggested that these normally have one driver at the wheel.

Other correlations show that if I wear a blue hat, I have more chance of seeing a blue car than if I wear any other hat. I've also found out that five days out of seven, white cars are the most common other than when I'm wearing a black hat on a Sunday. I also found out that yellow cars are most common when wearing a green hat but are the least common when wearing all other hats.

I've also been able to derive further factors due to these correlations. One is the Tuesday-red car significance which means that there is a high incidence of red cars appearing in low numbers on the following Thursday when wearing either a green/blue or white hat on a Tuesday.

There are a number of other factors which are subsequently 'loaded' to provide me with a means of predicting car colours on a particular day which I subsequently publish.

However, despite the science and validity of my factor analysis I find that my results are subsequently unreliable and meaningless.

There are a number of reasons for the unreliability of/ or invalid conclusions presented. They include:

1. Because I didn't start with a preconceived model, I failed to define what a car is for the purposes of my analysis. I subsequently included all vehicles save for certain heavy goods lorries, coaches, caravans etc. However, I did include white vans which I found out skewed my data from Monday to Friday. (Also by a methodological oversight I never wore a black hat on Saturday).
2. I had arbitrarily split the colour spectrum into six colours for the purposes of matching cars to hats. This is problematic as definition of greens, blues, reds and silver for example are open to interpretation (and eyesight). This over-simplified or skewed the data and, worse pretty much invalidated it.
3. I failed to take into account my own design bias which was that I don't like Mondays (ended up sampling rather than counting all) and green cars (many were given a bluish tint) which affected my results.
4. I failed to account for externalities that made some of my findings either naturally or un-naturally occurring by not defining upfront (white vans a particular case in point).
5. There was no reason to choose the day of the week other than for some form of arbitrary time period yet I didn't define if it was 24/7.
6. The relationship between hats and cars is completely spurious (other than for any externalities not taken into account).
7. By collating all types of incidental data I was able to find correlations on what turned out to be further spurious information, people wearing glasses, pets, furry dice, registrations, trailers, types of car etc.
8. Repeating the experiments would improve some of my statistical tests but then that's what happens if I keep collating data.

This is a simplified fictional scenario using 'factual' data merely to show the various dangers of insufficient preparation and analysis and how factor analysis can give lie to potentially irrelevant insight.

Qualitative assessments such as employee engagement rely largely on the robust design of the construct (i.e. its underlying model/hypotheses) as well as its survey instrument otherwise data will undoubtedly become spurious, insight meaningless and my whole experiment will suffer from the 'Garbage-in, garbage out' syndrome.

A review of the core process

Sampling

Best Companies use a random sampling process to collate statistically representative data and this follows standard statistical protocol for research purposes (circa 20%, excluding those with 50 employees).

However, there is a wider issue regarding sampling versus total population available for empirical collation, analysis and any subsequent organisational intervention.

We must remember that sampling is still indicative and care must be taken if we are to infer actual scoring that leads to organisation intervention. There is plenty of debate with regard to the two arguments. From a due diligence perspective we would always look to survey at least 50%, to be on the safe side pending any potential issues of representation.

Anonymity

Best Companies comment specifically on the anonymity of respondent issue. This is perhaps heightened given the sampling process.

However, claims of outright anonymity are false for three reasons:

- The random selected sample is e-mailed. Of course companies can easily find who has been e-mailed (whether advertently or inadvertently).
- Where e-mail cannot be used, the random selected sample is sent paper versions which are handled via a nominated internal project manager at the organisation.
- Where organisations have distributed locations, a number are pre-selected

Though Best Companies claims checks are in place to protect against potential undue influence this does effectively acknowledge a lack of anonymity and thus the process is open to undue influence.

There is possible debate over the overall level to which a group of employees can freely respond anonymously. Best Companies quite rightly surveys all organisations of 50+ (up to 250) employees but whether this is sufficient to overcome potential respondent social desirability

and/or acquiescence/fear over say, questions of leadership, is open to challenge particularly where subjective comment is also provided and used in any subsequent feedback.



Gaining Best Companies accreditation

The accreditation methodology uses scores from the core question set (BC16). They are also used for the 'Dip-Check' survey.

Given the evidence already provided in this evaluation we would reiterate our comment that due to the unreliable design of the core questions, we can place no reliance on the results provided for accreditation.

The Sunday Times-Best Companies Rankings

Since the inception of the Rankings in 2001, we observe that there is an annual 'churn' of around 50% meaning that there is effectively a new index every two years.

Yr	2002	2003	2004	2005	2006	2007
% churn	61	52	61	56	52	49

Coupled with the fact that the organisations contained within the published Rankings represent less than 0.7%¹⁰ of organisations in the UK it is hard to provide any credence or meaning to the Rankings themselves.

Even allowing for total submissions the figure is still only around 2%.

The concept of some form of any top100 in organisation performance, based on random submission and non-reporting for those that don't make the cut is quite seductive. However this form of ranking system is meaningful only if the following conditions are met:

- All participating organisations are reported to reflect (index or league table) the overall mix of constituents which provides a qualitative check
- That the submissions reflect a representative sample of the particular population (this can range from 20% upwards dependent upon the 'market')
- Alternatively, this system can apply to any existing published index which has credibility

The reality is that these indices can be too often seen as gimmicky due to their lack of representation. The current ST-BC Rankings would be viewed in this category.

Is the monicker 'best company' an accurate reference term?

The term 'best' implies a certain association which implies the 'best of' a certain group in this case companies/organisations. Normally with a definition of this sort, a defined criteria in relation to an index or classification is helpful to provide a comparative benchmark.

Also, the term 'best companies to work for' is potentially misleading as much of the rankings are based on benefits and as we have highlighted a flawed employee engagement construct.

Thus, this leaves open the potential to legal challenge, given indices, based on more de facto criteria which evaluate performance that includes employee engagement, or indeed organisations that have a robust employee engagement construct.

There is also the possibility that potential employees could be misled by the rankings if they are seeking employment.

¹⁰ DTI SME stats 2005
<http://www.dti.gov.uk/bbf/small-business/research-and-statistics/statistics/page38573.html>

Benefits to organisations participating in the Rankings

There is no question that in the absence of any meaningful differentiation in the SME market, there are a number of issues to raise regarding both successful and unsuccessful participation (i.e. ranking within the top 100 and no ranking at all).

We acknowledge that the ST-BC Rankings may have raised the profile of certain people management issues particularly for SMEs (notwithstanding the existing contribution of Investors in People). Certainly inclusion within the Rankings provides a higher profile of the organisation to the newspaper readership which amounts to (partly paid) advertising.

However, we would question as to the actual drivers for Rankings involvement and whether they are truly people-centric as opposed to recruitment marketing.

Certainly, organisations within the Rankings prominently feature the 'badge' on their websites, though we note in most cases that this only gives the year not the actual ranking itself (unless mentioned in accompanying text).

We have also found evidence that some organisations continue to promote their award from a previous year which surely defeats the object of the Ranking exercise. Technically, the Ranking should only be in existence for the current named year given the nature of the award as this could be deemed misrepresentative from a trading perspective or that the award itself becomes even more meaningless.

There is very little evidence available as to the benefits gained by those organisations appearing in the Rankings, save for their public exposure. Given the 'recruitment branding purpose' with which the Rankings were originally published, we would question as to the local benefits any small companies receive (who make up a large proportion of entrants).

It should be noted here that the Scheme is not open to organisation with less than 50 employees and less than 5 years trading thus bringing into question the actual SME focus.

Issues for organisations participating in the Rankings

Though an individual organisation appears to have some form of raised profile to potential employees by appearing in the ST-BC lists, the question is what benefits are gained?

As already stated there is no available research that has assessed the benefits despite public funding which has been assigned to the Rankings.

Many SMEs are based in single locations – thus for many potential jobs they will be unable to take advantage of national coverage in many cases. Does the national profile influence locally?

The dangers are in fact that these companies may in fact suffer from increased unsolicited CV traffic which may prove costly from an administrative perspective.

The particular marketing focus on associated work benefits may also be concerning. This may not lead to the hiring of highly engaged employees, but those who are merely seeking 'continuance commitment'.

The other issue to consider is whether or not an appearance within these lists, given the 'authenticity' provided by The Sunday Times and DTI, is associated with product-service quality and therefore provides 'trading' advantages¹¹.

If this is the case, then given the shortcomings of the underlying survey construct, the accreditation and the limitations of the 'Best100' as a credible index, this leaves the possibility of potential technical challenges under trading standards.

¹¹ For example there appears to be an over-representation of recruitment agencies given the normal sector distribution, i.e. 15% of small companies and 11% of medium-sized are recruitment agencies which would appear to support this surmise.



Best Companies - funding

Available data suggests that Best Companies originated through looking to replicate the 'American Best Companies to Work for' list.

As a result Best Companies (then branded 'Great Place to Work UK Ltd') received funding of £370,000 from the DTI in 1999.

*"DTI supports the research for this annual list and allocated £370,000 over three years in 1999. The DTI renewed its funding for the research for the 2003 list and is exploring options for 2004."*¹²

DTI representatives have since been quoted¹³ as seeing the relationship with Best Companies as beneficial, drawing attention to the robustness of the survey process and publicity benefits.

Despite this endorsement and involvement of DTI senior ministers with the awards¹⁴, the DTI does not appear to actively publicise its relationship with Best Companies beyond its inclusion in the Sunday Times published supplement.

Final comment

Though one can understand the opportunities and/or benefits of the ST-BC Rankings, the fundamental design flaws with which this review has highlighted does raise serious issues.

Given the evidence provided by this review, we would advise extreme caution to organisations and sponsoring organisations with regard to the serious credibility issues of the survey instrument, the ensuing 'engagement' framework, accreditation and Rankings.

There is also a potential issue of misrepresentation, both in terms of the meaning of 'best' in this context and to the potential misrepresentation of the employer brand to a prospective employee, should the construct be used as 'evidence' in recruiting.

The endorsement of The Sunday Times and DTI to the rankings (and by association any related interaction with organisations) unfortunately provides a false air of authenticity, and raises the questions around such endorsements.

While we appreciate the focus and interest on people management of this nature and its heightened effect in general, any potential flawed product will almost certainly be counter-productive.

We can understand the interest a national paper could have in this type of popular activity, after all newspapers are in the business of selling newspapers. However, we would not expect journalists to have technical knowledge of human capital management nor understand the vagaries and limitations of 'so-called' academic rigour from a statistical context, and therefore potentially exposing themselves to a less than robust construct.

Similarly, we would also acknowledge the DTI's interest from the SME angle and the potential benefits of raising any people profile from their perspective given the backing of project funding. However, we would advise that providing a perceived Government backing to what is potentially a flawed product can have repercussions of misrepresentation to potential users, not least from those firms who provide higher quality services in the field.

¹² House of Commons Hansard Written Answers for 17 March 2003, Column 521W, relating to Sunday Times 50 Best Companies to Work for 2001 and 100 Best Companies to Work for 2002.

¹³ Sunday Times (03/11/06)

¹⁴ See Rt. Hon. Patricia Hewitt's speech made at "100 Best Companies to Work For" Award Ceremony, London, Thursday February 27, 2003.

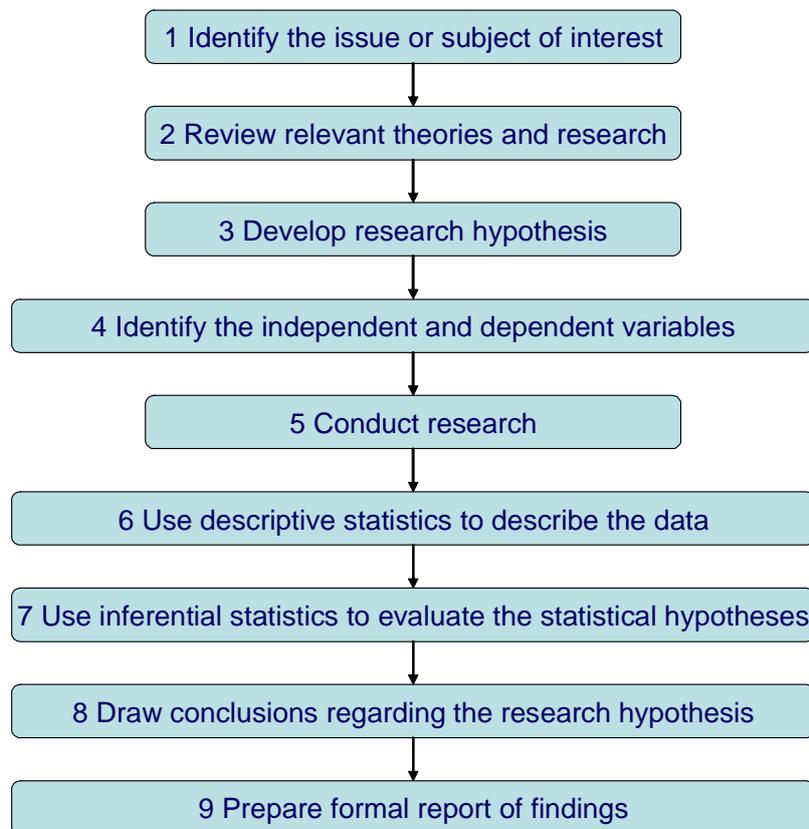
Bibliography

- Aiken L S & West S G (1991) 'Multiple Regression: Testing and Interpreting Interactions', Sage
- Aron A, Aron E & Coups E J (2006) 'Statistics for psychology', Pearson International 4th Ed
- Babbie E (2007) 'The Practice of Social Research', Thomson Wadsworth 11th Ed
- Child D (2006) 'The Essentials of Factor Analysis', Continuum 3rd Ed
- Cohen J, Cohen P, West S G & Aiken L (2003) 'Applied Multiple Regression/ Correlation Analysis for the Behavioural Sciences', LEA 3rd Ed
- Cohen J (1988) 'Statistical Power Analyses for the Behavioural Sciences', LEA 2nd Ed
- Cole D A (1987) 'Utility of Confirmatory Factor Analysis in Test Validation Research', Journal of Consulting and Clinical Psychology Vol 55 No 4 pp584-594
- Comrey A L (1978) 'Common methodological problems in Factor Analytic Studies', Journal of Consulting and Clinical Psychology Vol 46 No 4 pp648-659
- Cronbach L J (1951) 'Coefficient Alpha and the Internal Structure of Tests', Psychometrika Vol 16 No 3 pp297-334
- Crowley S L & Xitao F (1997) 'Structural Equation Modeling: Basic Concepts and Applications in Personality Assessment Research', Journal of Personality Assessment 1997 68(3) pp 508-531
- Cudeck R & O'Dell L L (1994) 'Applications of Standard Error Estimates in Unrestricted Factor Analysis: Significance Tests for Factor Loadings and Correlations', Psychological Bulletin Vol 115 No 3 pp475-487
- Darlington R, 'Factor Analysis'
<http://www.psych.cornell.edu/Darlington/factor.htm>
- Ford J K, MacCallum R C & Tait M (1986) 'The Application of Exploratory Factor Analysis in Applied Psychology: A Critical Review and Analysis', Personnel Psychology 39
- Fowler Jr, F J (2002) 'Survey Research Methods', Sage
- Fowler Jr, F J (1995) 'Improving Survey Questions', Sage
- Gorusch R L (1997) 'Exploratory Factor Analysis: Its Role in Item Analysis', Journal of Personality Assessment 68(3) pp532-560
- Haig B D (2005) 'Exploratory Factor Analysis, Theory Generation, and Scientific Method', Multivariate Behavioral Research 40(3) pp303-329
- Hair Jr J F, Anderson R E, Tatham R L & Black W C (1998) 'Multivariate Data Analysis', Prentice Hall 5TH Int Ed
- Idaszak J R, Bottom W P & Drasgow F (1988) 'A Test of the Measurement Equivalence of the Revised Job Diagnostic Survey: Past Problems and Current Solutions', Journal of Applied Psychology Vol 73 No 4 pp647-656
- Keppel G, Saufley Jr, W H & Tokunaga H (2002) 'Introduction to Design & Analysis', Freeman 2nd Ed.
- Kline P (1994) 'An Easy Guide to Factor Analysis', Routledge
- Kline R (2005) 'Principles and Practice of Structural Equation Modelling', Guilford 2nd Ed
- Lattin J, Carroll J Douglas & Green P E (2003) 'Analysing multivariate data', Thomson Brooks Cole
- Maydeau-Olivares A & McArdle J (2005) 'Contemporary Psychometrics (multivariate applications series)' LEA
- Mitchell M L & Jolley J M (2004) 'Research Design Explained', Thomson Wadsworth 5th Ed
- Monge P R & Cappella J N (1980) 'Multivariate Techniques in Human Communication Research', Academic Press
- Mulaik S (1987) 'A Brief History of the Philosophical Foundations of Exploratory Factor Analysis', Multivariate Behavioral Research 22 pp267-305
- Oppenheim A N (1992) 'Questionnaire Design, Interviewing and Attitude Measurement', Continuum
- Reyment R & Joreskog K G (1996) 'Applied Factor Analysis in the Natural Sciences', Cambridge
- Salkind N (2007) 'Encyclopedia of Measurement and Statistics Volume 1, Sage
- Shaughnessy J, Zechmeister E B & Zechmeister J S (2006) 'Research Methods in Psychology', McGrawHill 7TH Ed
- Thomson B (2004) 'Exploratory and Confirmatory Factor Analysis', APA
- Thurstone LL (1940) 'Current Issues in Factor Analysis', Psychological Bulletin Vol 37 No 4 pp189-236
- Velicer W F & Fava J L (1998) 'Effects of Variable and Subject Sampling on Factor Pattern Recovery', Psychological Methods Vol 3 No 2 pp231-251
- Wirth R J & Edwards M C (2007) 'Item Factor Analysis: Current Approaches and Future Directions', Psychological Methods Vol 12 No 1 pp58-79



Appendix

Conducting scientific research



Source: Introduction to Design and Analysis, Keppel G, Saufley W Jr, Tokunaga H, 2nd Ed 2002

[Note: Commonly experiments are repeated by iterating between steps 3 through 9 allowing and refining hypotheses].

