SPE 61228

The Hearts and Minds Project in an Operating Company: Developing Tools to Measure Cultural Factors

P.T.W. Hudson, SPE, Leiden University and F.C. Willekes, Petroleum Development Oman

This paper was prepared for presentation at the SPE International Conference on Health, Safety, and the Environment in Oil and Gas Exploration and Production held in Stavanger, Norway, 26–28 June 2000.

Abstract

The Hearts and Minds project is being carried out to improve safety performance across the Oman oil industry. The objective is to identify how people can become intrinsically motivated for good HSE behaviour and to develop (micro-) tools to achieve such well motivated behaviour in a simple and effective manner, optimally tuned to the cultures in the organisation. The Hearts and Minds programme involves two components. One is the development of tools to measure the safety culture, readiness to change and national cultural factors. The second involves the development of 'micro-tools' to produce development of safety culture. This paper reports the results of the first stage. Reliable and simple scales to measure safety culture and readiness to change have been developed. The results show differences between PDO and a number of its major contractors. Most interestingly there are reliable differences between managers, supervisors and workforce that are independent of the company. Managers are invariably more open to change and have higher evaluations of the safety culture than do either supervisors or the workforce.

A replication of Hofstede's scales for the measurement of national culture was also successful. The scales measured were Power Distance, Uncertainty Avoidance, Collectivism-Individualism and Masculinity-Femininity. Significant differences have been found between different groups that have considerable implications for the development of safety culture and the implementation of initiatives in general. This paper also discusses the consequences of the results for implementation of safety initiatives within a company that is characterised by being culturally extremely diverse.

Introduction

One of the greatest problems facing a multi-national organisation, operating in the global market, is the diversity of cultures with which it has to work. In times gone past it was sufficient to stamp a single corporate culture, emanating from head office and propagated by expatriate managers. Today the roles are reversed, cultures have to be respected and the expatriates are now in head office. Furthermore the realisation has grown that one of the reasons why management, such as the implementation of safety programs, has been difficult is because those programs were forced upon the work force, top-down, in a 'one size fits all' methodology. In early days this approach worked and safety performance improved markedly. But this is not entirely surprising as almost anything was likely to have some positive effect. More recently the systematic safety management approach, generalised to HSE-management systems, has been predominantly culture free. The main requirement for acceptance of HSE management systems within an organisation is the possession of an organisational culture that is comfortable with bureaucratic control.

The next stages, however, have become increasingly difficult. Improvements in performance have been less forthcoming. The HSE tools project[1] identified two specific types of culture that needed to be taken into account. These are (i) the organisational culture, how does the organisation think and react, and (ii) the national or geographical culture of the workforce, how do the people think and react? A new approach to managing safety needs to be compatible with the current stage of development. Tools have to operate in ways consistent with the national characteristics. The aviation industry has learned this lesson the hard way[2]. The original introduction of Crew Resource Management, which became an internationally mandatory requirement set by ICAO, floundered initially because of the assumption that the commercial aviation culture was homogenous. In fact the assumption was that the cockpit was a 'culture free zone' so no account need be taken of differences in the way different nationalities respond to situations, events and other people. This assumption turned out to be very wrong[3]. For instance, different cultures turn out to take quite distinct stances with respect to cockpit automation, a feature that, at first glance, would seem to be culture free. Recently CRM training

programs have been developed which take much more account of national cultural differences.

Measurement Tools for Cultural Factors

If cultural factors are so important, what can you do? The primary requirements are to be able to measure such 'soft' factors as culture and, using the results of such measurements, to be able to develop tools that are optimally effective in specific cultures. How can you measure cultural factors in the field? The Hearts and Minds Program[1] is intended to develop an intrinsically motivated workforce, which is interpreted as developing a safety culture. Within the Shell Group there is a wide range of different types of organisation in very different parts of the world. The expectation is that the differences that may be encountered have to be respected in order to achieve success. In order to do this it is necessary to measure the safety culture, the readiness to change of the workforce and the national characteristics of those involved. While it may be possible to use a 'one size fits all' approach in the early stages of development of an organisation, the later stages require a more customised approach if success is to be guaranteed and failure avoided.

The workforce in Petroleum Development Oman (PDO) and its contractors is highly varied and has reached a level of safety performance where the problems of cultural sensitivity are likely to become acute if further improvement is to be obtained. As well as the Omanis, who make up the bulk of the company and contractor workforce, there are European and American expatriates, Pakistanis and Indians from different parts of the sub-continent, and Filipinos as well as a number of other nationalities on cross-posting such as Nigerians and Malaysians. There is a considerable contractor workforce ranging from major multi-national drilling and exploration companies to more local engineering operations. PDO provides what is possible the most diverse cultural mix in Shell Group's world-wide operations.

Specific Tools to Measure Cultures

This paper reports on the development of three measurement tools for use in a wide variety of settings. All three tools were developed and calibrated at the same time.

1. The first test is the *Readiness for Change* scale. This can be used to assess whether the members of an organisation are open to change, or feel already actively involved in a change process or, feel that no improvement is possible or necessary. The latter may be seen as indicating complacency.

2. The second test is used to assess differences between *National Cultures.* The items are drawn from the test originally designed by Hofstede for use in IBM and have been used elsewhere in world-wide commercial aviation. Knowledge gained by using these scales can mean the difference between success and failure in working with culturally diverse groups.

- The third set is the Westrum *Safety Culture* scale. This is a test intended to measure the organisational safety culture. These are seen as evolving from the Pathological through the Calculative to the Generative. This test offers a quick way of measuring the culture people feel most appropriate to describe their organisation.

The tests are usually based upon a single sheet of A4. The Westrum scale was administered by making up plasticised cards, requiring respondents to choose one of five for each of the dimensions used in the test. Such tests are intended to be given anonymously to groups and the only information required is group membership (e.g. drilling, production etc. or Operator/Supervisor/Manager). Small samples are usually sufficient; it is not usually necessary to survey all the members of an organisation. It is not advisable to give any of these tests to individuals as they are designed to assess groups and are not tested with a view to their reliability with individuals.

This section reports the results of analysing the data collected in PDO and its contractors. There were a total of 366 people interviewed. The breakdown of the samples is give in Tables 1 and 2. The proportion of missing data, where an interviewee did not complete part of the questionnaire, is rarely above 10%. The separate items can be aggregated into scales, which are formed by averaging over two or more items. The specific items used are contained in Annex A, which is the questionnaire form as it was used. The five Safety Culture dimensions were placed upon cards so that for each dimension a set of five cards was presented to the interviewee.

In a number of cases the scales are reported with a measure of scale reliability called Cronbach's α. A value of α in excess of 0.7 is very good, values lower than 0.6 are weaker, but where there are few questions (less than 5) used to make a scale, α is not very meaningful. The scale values are commented upon where appropriate.

Company Type	Number of interviewees
PDO	108
Construction	129
Service companies	82
Drilling companies	47

Table 1. Numbers of individuals interviewed per company type

Job Category	Number of interviewees
Manager	49
Supervisors	140
Operator/technician	174
Unknown	3

Table 2. Numbers of individuals interviewed by job category

Results are generally reported in terms of company (See Table 1 for numbers of interviewees) where several contracting companies are pooled. Results are also reported in terms of Job Category, over Managers, Supervisors and Workforce. Table 2 shows the frequencies of interviewees in the different job categories.

The tests were administered, usually on paper but verbally when there were doubts about an interviewee's ability to read and write English. Interviews were carried out by a psychologist and generally took between 30 and 60 minutes. Where a translator was necessary they tended to take about 1 1/2 hours. Some interviews were held in languages that the interviewers also spoke (German and Arabic).

The Readiness to Change Test

The Readiness to Change test provides a way of measuring whether people are ready to improve their safety performance and culture or are comfortable where they currently are. Groups that score low may often be regarded as complacent. There are three items, one of which respondents have to choose for each of five dimensions: Communication, Organisational Attitudes, HSE, Organisational Attitudes and Working Behaviour. The items, for each dimension, ask whether:

1. Things are as good as it gets round here
2. Things could get better
3. We are working to improve round here

The test can be given to any level in the organisation. There should be at least 10 respondents per group and it is important to stress that the responses are anonymous. The number computed is the average "readiness to change" number of the group. A score between 1 and 1.5 means that people are pre-contemplative, meaning that people don't see a problem and no need for change. A score between 1.5 and 2.5 means that people are aware of the problems but they don't know how to solve it (yet). These people are contemplative. A score above 2.5 indicates that people are fully aware of the problems but more important, they are actually working on it to improve their situation. These people are in action. Scores below 2 strongly suggest a degree of complacency, unless an initiative has just finished, when people need to be left to get used to their new situation.

Results. The Readiness to Change scale has a Cronbach's a of 0.77, which is reliable. Significant differences were found over both companies and job categories. Figure 1 shows that while there are significant differences between companies, in all cases managers are more advanced than supervisors, while the workforce lags behind both.

A more detailed analysis of the transition scale results is revealing. In particular it is possible to regard the proportion who are Pre-contemplative as a measure of complacency, or even ignorance, within an organisation, while the proportion who choose Action is a measure of activity. Table III shows a degree of complacency at the workforce and supervisory levels when compared with managers.

Job Level	Pre-contemplative	Contemplative	Action
Operator/ Technician	36%	54%	10%
Supervisor	39%	47%	14%
Manager	15%	65%	20%

Table 3. Percentages, within job-levels, of individuals found in the different transition stages. What is noticeable is that managers are considerably less pre-contemplative than the rest of the workforce, where nearly 40% are satisfied with the current state of affairs, and are much more involved in action.

Company Type	Pre-contemplative	Contemplative	Action
Company	21%	68%	11%
Construction	45%	48%	7%
Drilling	36%	43%	21%
Service	27%	55%	18%

Table 4. The percentages of employees within different organisations found in the different transition stages.

In Table 4 there are strong results for complacency within the construction companies (45%) and a rather passive understanding, that there are still problems, within the Company (68%), but without associated action. The Drilling and Service organisations report considerably more activity, but there are still many complacent drillers, it is as if drillers are more likely to go to action once they realise that action is appropriate.

The transition measurement instrument appears to be simple, highly robust and ready to use. Its use enables those who wish to change an organisation to assess how much effort this is likely to involve. The use of the proportion of precontemplative staff as a measure of complacency is to be recommended as it appears to work and produce realistic results that can distinguish organisations.

Hofstede National Culture Scales

The Hofstede National Culture test provides a way of measuring National Culture of a company. The reason for measuring National Culture is that differences in for example Power Distance have to be taken into account when making optimal change tools. Knowledge about how national groups score on the National Culture scales can be used to avoid making implementation errors. The scale is made of eleven items measuring the following five cultural dimensions: Power Distance, Uncertainty Avoidance, Masculinity, Femininity and Collectivism-Individualism. See definitions of the dimensions in Interpretation section.

People at any level in the organisation can be tested. Significant differences have been found between different national and geographical cultures. It is recommended to have at least 10 individuals, preferably 10% of the total group and to stress that the results are anonymous. These tests need only be done once.

Uncertainty Avoidance is assessed by measuring agreement with two questions:

- It is OK to break the company's rules if it is good for the company.
- Sometimes I am not sure what to do at work, and this makes me feel nervous or tense.

A score above 3 on **Uncertainty Avoidance** (UA) indicates that the respondents have a *high* Uncertainty Avoidance level. This means that they prefer to work under strict conditions without any exceptions. A *low* score indicates that the people are willing to break the company rules if it is good for the company and that they feel they know what to do in undefined situations.

Power distance is assessed by requiring a frequency measure for two questions:

- How often at work are you afraid to disagree with your boss?
- How often at work does your boss ask for your opinion?

A score above 3 on **Power Distance** (PD) indicates that there is a flat and open organisation where people are not afraid to discuss with their superiors (Low Power Distance). The workers have the feeling that there is listened to them by the superiors. A score under 3 indicates a distinct hierarchical structure, where people are less comfortable to express their ideas and ask questions to their superiors (High Power Distance).

The remaining scales were determined by asking people to consider their ideal job (not their current one) and then to indicate how important each one is to them, using a scale from very important to very unimportant, for the following items:

1. Maintaining good interpersonal relationships with fellow workers or crewmembers
2. Having the opportunity to get promoted
3. Having security of employment
4. Living in an area desirable to you and your family
5. Having an opportunity for high earnings
6. Having enough time left for your personal or family life
7. Working with people who co-operate well with one another

- A score above 3 on **Masculinity** (questions 2 and 5) indicates a society in which gender roles are clearly distinct; men are supposed to be assertive, tough and focused on high achievement (high earnings, promotions etc).
- A high score (above 3) on **Femininity** (questions 1, 3 and 7) indicates a society in which there is a greater concern for quality of life issues such as quality time and living in a nice area.
- A high score on (above 3) **Collectivism-Individualism** (questions 4 and 6) indicates a Collectivist culture in which behaviour is linked to the in-group, extending beyond the immediate family. A low score indicates a Individualist culture that is characterised by consideration of behaviour in a narrowly defined area of personal costs and benefits.

Results

Uncertainty Avoidance. Uncertainty Avoidance is a single item scale after the one other scale item produced extremely skewed items, so the α value is meaningless. The item that was rejected asked whether respondents would deviate from procedures if it was in the organisation's interest. Hardly anyone felt they could respond to this, even if they do in practice. This means that the item is useless as a scale item, but the finding suggests that more work needs to be done in understanding whether violations really occur or that people are aware that there is a 'correct' answer. With the single remaining item there is a significant difference between companies, almost all due to a *stronger* uncertainty avoidance in the construction companies. However it must be noted that a mean of 2.0 implies personnel who have few problems with coping with uncertainty at work. Unsurprisingly drillers have fewest problems with uncertainty.

Power Distance. The Cronbach's α for Power distance was *0.34*, which is marginally acceptable for a scale with a small number of items (2). The values being below 2.5 means that all personnel feel comfortable that there is an atmosphere of interpersonal co-operation and mutual influence. There are no significant differences between companies, however the differences between all three job levels are significant on this scale. Figure 2 shows that managers felt the least power distance and operators and technicians somewhat more.

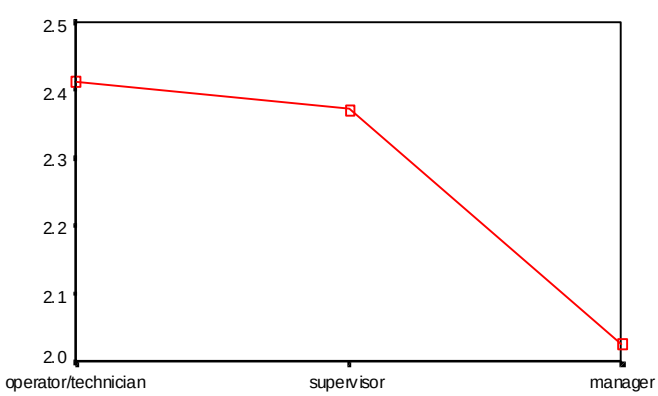

Figure 2. Power Distance measured over job levels.

Masculinity and Femininity. 'Masculinity' is mostly concerned with wanting to have the opportunity to make high earnings and be promoted. The scale α is *0.48*, which is

reasonable for a two-item scale). The effect of company is marginally significant ($p < 0.06$). Post-hoc comparisons between the company means show that only the means of Service and Construction differ significantly from one another ($p< .005$), with those in Construction less concerned with masculine work values than those in Service companies.

'Femininity' is concerned with having security and working with people who co-operate well together. The scale α is 0.53, which is reasonable for a two-item scale. The overall effect of company is significant ($p< 0.02$). A higher score on femininity (Service companies) means that people seek more security and want more time for their personal lives. The construction industries show lowest femininity scores. There are no effects of job level on femininity.

In general companies seem to have the same pattern of results for both masculinity and femininity, which suggests that those who score highly on both want to have their cake (promotions, money - masculine) and eat it (security, time with families - feminine). The construction workforce from the Subcontinent is possibly self-selected in putting up with poor wages and no families.

This dimension is probably the least useful, even if illuminating. It does, however, provide an indication of the types of rewards people find important in different national cultures. Interpretation of the interview questions about the ideal job is very susceptible to interpretation and, therefore, to the ability to understand English.

Individualism-Collectivism. This scale is particularly relevant for implementation purposes within organisations. Cronbach's α is *0.57*, which is acceptable with 3 items. On the whole all organisations showed a tendency to being collectivist, being happy to work in groups to a common goal. Companies differed very significantly, with Service industries most collective and construction companies most individual. There is a slight trend for managers to be more individualistic than the rest of the workforce.

Nationality. Nationality is, clearly, the most crucial factor in determining responses to the Hofstede scales and, therefore, in influencing the ways in which implementations might be affected. The following effects are all significant over differences in nationality: Power Distance, Masculinity, Femininity and Individualism-Collectivism. The Readiness to Change scale is also sensitive to differences in nationality.

Discussion. The Hofstede scales are effective in discriminating between different companies and between different job levels. The results here highlight the importance of taking account of such differences and, therefore, the importance of measuring to start with. The differences between managers and the rest of the workforce are especially important because managers will need to understand that the rest of their workforce does not necessarily think like them.

Westrum Safety Culture Scale

The Westrum Safety Culture test provides a way of measuring the Safety Culture of an organisation. The reason for measuring Safety Culture is to find out how people are actually thinking about safety and how safely they work. Differences between the different stages of safety culture have to be taken into account when designing optimal HSE tools and when trying to improve performance. It is crucial to know where the organisation is if you want to change it.

Safety Culture is measured using a range of different cultures. These vary from

- **Pathological** - There are few if any safety rules, the workforce is supposed to look after itself and is responsible for accidents. Accidents are part of the job.
- **Reactive** - Safety starts to be taken seriously, but only after incidents is there any action.
- **Calculative** - Safety is treated very seriously, there are management systems and much collection of data
- **Proactive** - People try to avoid accidents and start to take a more bottom-up approach
- **Generative** - All levels participate in active search for data, there is chronic unease coupled with exemplary safety performance

This scale is made of 5 items measuring the Safety Culture as defined above along each of the following five dimensions:

1. **A** -Communication
2. **B** - Organisational Attitudes
3. **C** - HSE
4. **D** - Organisational Behaviour
5. **E** - Working Behaviour

Score	Interpretation of Results	Actions for Improvement to Next Stage in Safety Culture
1 – 1. 5	Pathological	Ensure that safety is taken seriously, develop awareness that things can be different
1.5 – 2. 5	Reactive	Start to develop systematic approaches to HSE Management
2.5 – 3. 5	Calculative	Realise that HSE management involves managing the unexpected
3.5 – 4.5	Proactive	Bring everyone together, break down barriers, integrate HSE into the business
4.5+	Generative	Keep working and avoid backsliding because of success in HSE performance

Table 5. *Scoring method for the Westrum Safety Culture Scale*.

Interviewees chose the most appropriate description for their organisation from a set of five plasticised cards for each dimension. Cronbach's α for the Safety Culture Scale (also called The Westrum Scale) is 0.73, which is very good. The differences between company and contractors and across job types were not statistically significant, but were indicative. The failure to reach statistical significance of the difference between organisations may well be due to the homogeneity of the different companies' safety cultures, driven by the company's strong influence. The strong α suggests this interpretation, but this hypothesis can only be tested elsewhere in the world, or in organisations that are clearly different.

Table 6 Shows the breakdown in choices made over the three job levels. Here it is clear that the modal (most frequent) choice is for proactive, but that managers are much more likely to characterise the company as Calculative than are the remainder of the workforce, possibly because it is they who are doing the calculating and managing the management systems. Interestingly managers are also less likely to choose for Reactive than are the others.

Conclusion

This paper has presented three tests of cultural values that can be used to discriminate between companies and between different levels in the work force. All of the tests are acceptably reliable and have generated interesting results. In general it can be stated that implementation of new HSE programs without using the knowledge such tests can provide is liable to be considerably more difficult and less likely to succeed than when such information is collected and used.

The results already obtained highlight a specific problem that needs to be addressed, and probably anywhere in the world. Managers regularly and reliably differ in their scores from other members of the workforce. They are more ready to change, have less power distance and uncertainty avoidance and are more individual. They are also more sanguine about how far their organisations' safety cultures have developed. Taken together this means that managers pressing for improvements may well attribute their own characteristics to those they manage, incorrectly as we have shown. Finding such incompatibilities may provide an explanation for why so many initiatives that seem so reasonable to management nevertheless fail.

What this paper has shown is that it is feasible to make tests of 'soft' issues that can be at least as meaningful and reliable as those used to assess 'harder' issues, such as sub-surface characteristics. What this paper also shows is that it is important that tests, often framed as questionnaires, need to be scrutinised and subjected to professional analysis. Many questionnaires, for instance, are used to assess workforce attitudes without any rigorous backing to the items used other than face validity. What is more there are many techniques for extracting more information out of such questionnaires than is immediately obvious.

References

1. Hudson, P.T.W. & van der Graaf, G.C. (2000) HSE Tools: Which are Appropriate? This conference.
2. Maurino, D.E. (1996) Eighteen years of the CRM wars: A report from headquarters. In B.J.Hayward & A.R.Lowe (Eds.) *Applied Aviation Psychology: Achievement, Change and Challenge*. Avebury: Aldershot.
3. Meshkati, N. (1994) Cross-cultural issues in the transfer of technology: Implications for aviation safety. *Report of the International Civil Aviation Organisation Flight Safety and Human Factors Regional Seminar and workshop, Amsterdam.* ICAO: Montreal.
4. Hofstede, G. (1991) *Cultures and organisations: Software of the mind.* McGraw-Hill: Maidenhead, UK.
5. Westrum, R.J. (1991) Cultures with Requisite Imagination. In J. Wise, P. Stager & J. Hopkin (Eds.) *Verification and Validation in Complex Man-Machine Systems.* Springer, New York.

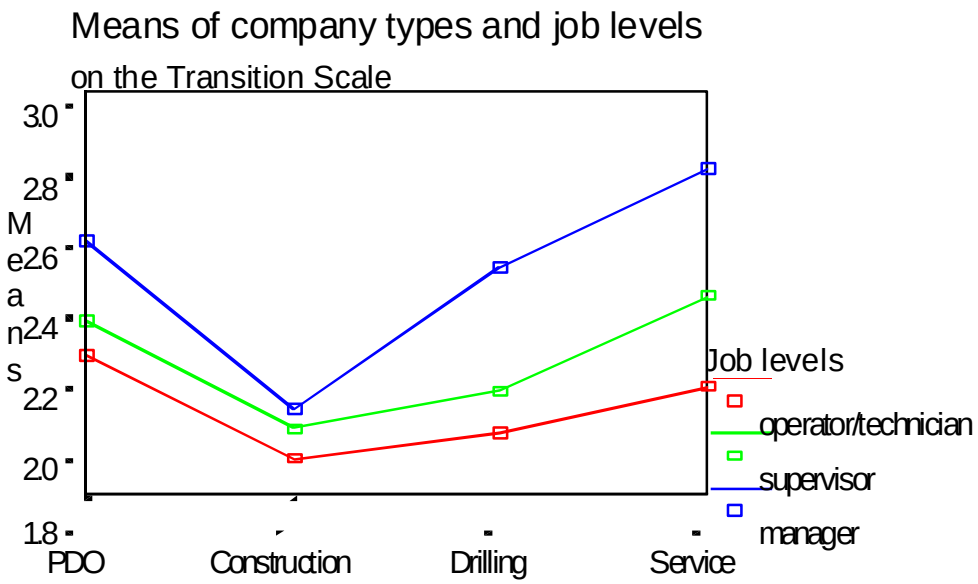

Figure 1. The Transition scale values for companies and job categories. The low values (< 2.0) represent a lower readiness to change – "Things are good round here". Managers are generally more ready to change, the workforce is least ready.

Job Level	Pathological	Reactive	Calculative	Proactive	Generative
Operator/ Technician	2%	12%	26%	59%	1%
Supervisor	1%	12%	28%	59%	0%
Manager	0%	7%	44%	49%	0%

Table 6. The scores on the Safety Culture Scale for the different job levels. The most striking results here are the similarity of operators and supervisors and the more conservative evaluations of the managers. The suggestion is that managers are better calibrated.Even so the modal answer of all groups is proactive.