DEVELOPING A BENCHMARKING SERVICE FOR THE HSL SAFETY CLIMATE TOOL†

Nicola Healey[1], Caroline Sugden[1], Sarah Binch[1], Chrysanthi Lekka[1] David Bottomley[2]
[1]Health and Safety Laboratory, Buxton, UK
[2]Merlin Analytical Service, UK

Since the launch of the Health and Safety Laboratory's (HSL) Safety Climate Tool in 2010, over 40 companies from across a range of sizes, sectors and locations around the globe have purchased the tool to use within their organisations. The Safety Climate Tool measures perceptions of health and safety within an organisation, and has 40 questions and 8 factors (see http://www.hsl.gov.uk/health-and-safety-products/safety-climate-tool.aspx). Current users and potential customers have expressed an interest in being able to compare their results on the tool with others. In response to this demand, Human Factors specialists at HSL have been working on developing a benchmarking service to enable them to do this.

As part of developing a benchmarking service for the HSL's SCT, a number of considerations had to be addressed and will be described in this paper:

- How much data is sufficient to offer benchmarking? Statistical questions around variation will be summarised.
- Do sector differences matter? For example, can lessons from the Olympic Park on safety culture inform the process industry?
- Are there across country-differences in safety climate data, and what are the implications of this for global operations?
- How can benchmarking, as a process, help organisations achieve a more positive safety culture?

The challenges of benchmarking and the implications from the data collected will be explored and discussed.

KEYWORDS: Safety Climate Tool, Benchmarking, Performance

1. INTRODUCTION

The Health and Safety Laboratory (HSL) revised the original Health and Safety Executive Climate Survey tool (HSCST) – now known as the HSL Safety Climate tool (SCT). The original HSE tool was used by approximately 825 organisations, many of which also signed up to a benchmarking service, run in conjunction with the Opinion Research Corporation International (ORC).

Revision of the SCT resulted in a reduction in the number of statements and slight modifications to the tool structure. As a result, the previous benchmarking data was no longer usable.

Users of the HSL SCT expressed a great deal of interest in benchmarking, and some initial consideration was given to the appropriateness of transposing the old data onto the new factor structure. Initial attempts to create benchmarks using a similar methodology to the HSE Stress Management Standards has also been attempted (e.g. collaborative work with the Olympic Delivery Authority), however more rigorous investigation regarding the most appropriate methodology for developing benchmarks was required.

Therefore, the objective of ongoing work by HSL was, building on initial considerations regarding a benchmarking methodology, to identify the most appropriate methodology for calculating benchmarking data for the HSL Safety Climate Tool, and to consider the practical implications associated with developing a benchmarking capability for users of the SCT.

1.1 SAFETY CULTURE AND THE SAFETY CLIMATE TOOL

Safety climate as a term was initially used by Zohar (1980) to describe attitudes towards safety, and was derived from earlier work on organisational climates (Tagiuri and Litwin, 1969). Safety climate tends to be explored through questionnaires exploring attitudes and perceptions regarding safety; and is a statistical construction of perceptions held in an organisation regarding safety (Rousseau, 1988). Cox and Flin (1998) commented that the terms safety culture and safety climate are often used interchangeably to refer to similar concepts. Culture, as a topic is viewed as much harder to study and measure (e.g. Schein, 1990), whilst climate is considered to be easier and can provide an indication of people's perception and cognitive interpretation of organisational practices (Lee, 1981).

The International Atomic Energy Authority used the phrase safety culture to describe the issues surrounding the Chernobyl incident (IAEA, 1986). There was a realisation that organisational structure (i.e. the roles and their

relationships, rules and procedures) was limited in achieving good health and safety performance. In fact, HSE stated that the explicit and implicit goal of a safety management system should be the development of a positive safety culture (HSE, 1991). This reflects an understanding that safety management systems succeed if individuals are motivated to comply and conform with the organisation's systems, which is where the need to understand the impact of safety culture arose. HSC (1993) defined safety culture as:

> "the product of individual and group values, attitudes, perceptions, competencies, and patterns of behaviour that determine commitment to, and the style and proficiency of, an organisation's health and safety management. Organisations with a positive safety culture are characterised by communications founded on mutual trust, by shared perceptions of the importance of safety and by confidence in the efficacy of preventive measures" (p. 23).

The development of a good safety culture is recognised as central to achieving good health and safety within an organisation.

The Safety Climate Tool is a survey tool developed by researchers at HSL and is based on a tool developed and sold by HSE in the 1990s. It has 40 items that cover 8 factors (see Figure 1). A full description of the revision of the SCT is provided in Sugden et al. (2009). Participants respond to the SCT questionnaire using a 5-point Likert scale, with responses ranging from 'Strongly Agree' to 'Strongly Disagree'. Each statement in the SCT questionnaire relates to one of a series of factors, and responses are used to calculate a mean score for each factor. The responses can be interpreted such that a good safety culture has a mean score closer to 5, whilst a poor culture has a mean score closer to 1.

1.2 THE SCT AND BENCHMARKING

Benchmarking can be defined as "a method of measuring and improving ... organisational performance by comparing ... with the best" (Stapenhurst, 2009 p. 6). Benchmarking organisational performance in the context of safety culture assessment is achieved through comparison of safety climate scores with those of other organisations' in order to identify, adopt and adapt practices to help bring about improvements in safety climate.

Benchmarking can be employed as a one-off 'audit' activity, or as part of a commitment to continuous improvement. Benchmarking performance against those of other companies enables an organisation to identify areas where they are doing well, as well as identifying areas to target for improvement.

HSL carried out market research with organisations that had either piloted or purchased and used the SCT. Some of the organisations within the market research sample had used the original HSCST, and the associated benchmarking service. Organisations who have used the SCT are commonly interested in how they compare to other organisations, and most participants had clear requirements from a HSL benchmarking capability:

- The ability to carry out Industry/sector-based benchmarking;
- The ability to carry out benchmarking analyses based on organisational size; and
- An interest in sharing good practice.

Benchmarking SCT performance enables organisations to:

- Quantify their current Organisational safety culture (within the context of other organisations' performance);
- Identify the gap between their organisation's SCT score and other organisations' SCT scores (e.g. the 'top performers');
- Quantify the potential improvement/possible gain for the organisation to operate at the level of the top performer.

There has been strong interest in a benchmarking capability to complement the HSL SCT from users of the original HSCST and the revised HSL SCT. Current users of the tool also have a number of specific requirements from a benchmarking service, and therefore the objectives of a HSL SCT benchmarking capability are to enable:

1. Comparison of performance to identify organisations achieving a positive safety culture;
2. Improvements and adoption of practices leading to a positive safety culture.

2. METHOD

To date, a number of approaches to benchmarking the SCT have been explored. Figure 2 provides an overview of the process that was followed, which is discussed in more detail in the following sections.

3. RESULTS AND DISCUSSION

3.1 LITERATURE REVIEW

The objective of the literature review was to identify best practice in collecting, analysing and utilising benchmarking data, specifically in relation to attitudinal survey data. The initial search resulted in over 250 results, of differing levels of relevance.

Despite some papers discussing benchmarking in specific and highly technical contexts, there was a distinct lack of detail regarding the benchmarking methodology employed. Throughout the literature search it was also difficult to distinguish between best practices in benchmarking (i.e. the best approach for benchmarking specific measures), as opposed to the benchmarking of best practices (e.g. comparison of best practices in a specific context across organisations).

A small number of studies that used benchmarking were identified from the literature review; however in

Figure 1. Safety Climate Tool factor structure

most cases these articles contained limited information, making it difficult to decipher the precise benchmarking approach that was used (e.g. Heidegger et al., 2002; Sexton et al., 2006). Mearns et al. (2001) noted that 'published accounts of benchmarking specific to the area of health and safety management are limited' (p. 772), and more than a decade on, this still appears to be the case. There was only one comprehensive study (Mearns et al., 2001) that benchmarked safety climate across organisations.

A review of the limited available literature showed that there are different approaches to calculating benchmarks. The use of these approaches tends to be data-driven (i.e. use of an approach that bests fits the data), rather than representing best practice. Stapenhurst (2009) outlined the different steps involved in a typical benchmarking analysis. The first step includes an initial assessment of performance levels within a given sample using some form of pictorial representation, such as a histogram. This allows an initial assessment of the range of performance levels across the sample, which are typically ranked from high to low. The second step in a benchmarking analysis involves comparing participants' performance levels against a target performance level or benchmark. Drawing on Stapenhurst (2009), there are three main approaches to calculating a benchmark each with its respective advantages and disadvantages:

1. Use of quartiles. This approach involves slicing the data into quartiles allowing participants to compare their performance levels against the top 25% of the sample. The use of quartiles as an indicator of target or good performance levels is intuitively appealing for organisations and can be a robust metric, as it is unlikely to be influenced by the values of a small number of participants. However, quartiles may often result in artificial distinctions between high and poor performers (i.e. as participants with the same scores can often be grouped in different categories depending on the sample size) or may mask meaningful differences between participant performance levels by grouping them within the same quartile.
2. Use of averages. Another metric that can be used as a benchmark is the numerical mean score. For example, an organisation's performance may be compared with an overall average score of all participants/organisations in a given sample. Stapenhurst (2009) argues that, as was the case with the top quartile, the use of the average or mean score is intuitively appealing because most organisations want to know whether their performance is above or below average. However, comparing a participant's performance against an average target performance is unlikely to be appropriate for all participants, or indeed meaningful depending on the nature of what is benchmarked.
3. Identifying the 'best in class'. This approach involves comparing a participant's performance against the best performer within a given dataset. The disadvantage of this approach is that the 'best in class' may represent an unrealistic target performance level or may be an inaccurate representation of performance; for this reason quartiles are often preferred as they are more likely to represent a realistic and 'achievable' benchmark.

The small number of relevant studies identified appeared to use one, or a combination of the approaches outlined above. In one of the most comprehensive benchmarking studies, Mearns et al. (2001) compared the safety

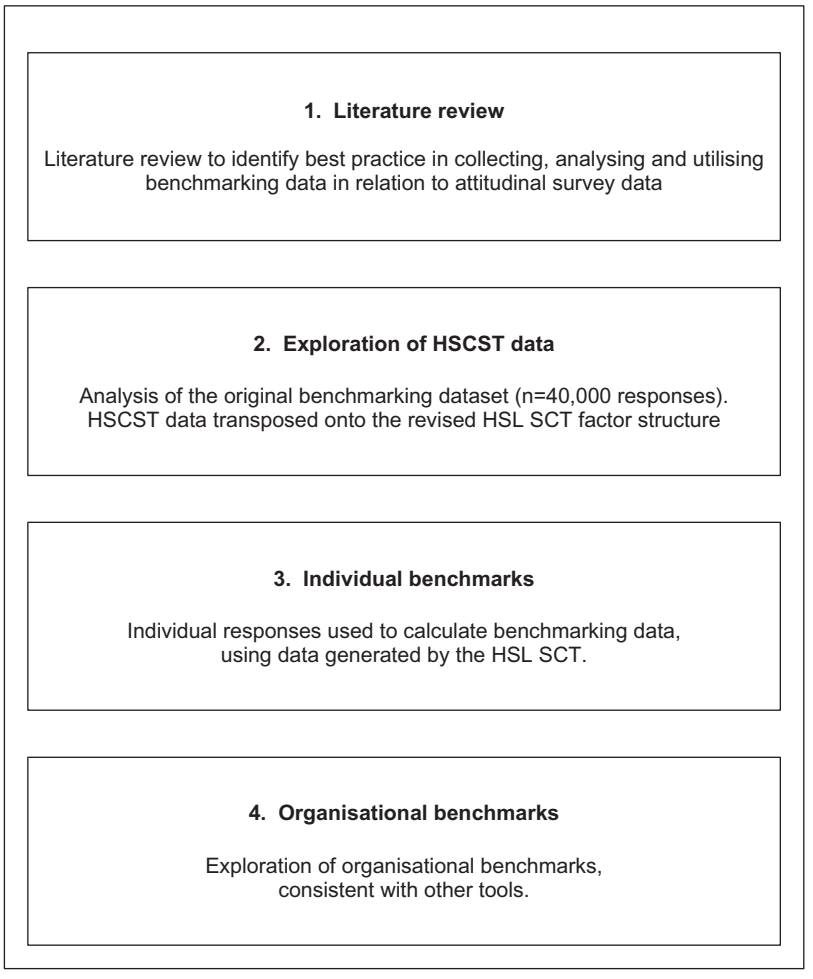

Figure 2. Overview of Benchmarking development process

climate profiles of nine oil and gas companies using the Offshore Safety Questionnaire (OSQ). Their measure captured six dimensions of safety climate: satisfaction with safety activities, perceived supervisor competence, perceived management commitment to safety, willingness to report incidents, frequency of general unsafe behaviour, and frequency of unsafe behaviour under incentives (i.e. unsafe acts due to peer group or management pressure). Their approach to benchmarking involved calculating mean scores for each of the six OSQ factors, and comparing the performance of each company with its peers by, (i) plotting performance on each factor using the mean and standard error of the mean, and (ii) using a statistical test to corroborate the observed differences in mean scores using graphs. In other words, their benchmarking approach was two-fold; the factor means for each company were first compared visually, using a histogram, and statistically in order to ascertain whether the observed differences were in fact significant.

Miller and Cox (1997) conducted a postal survey in order to benchmark six sectors (e.g. chemical and allied industries, food and drink, government, manufacturing, oil and gas and transport) in terms of their compliance with the Health and Safety Executive's (HSE) health and safety management principles. In particular, they compared responses to each of six health and safety management principles, in line with HSG 65 (1997) 'Successful Health and Safety Management': initial and periodic status review, policy and objectives, organising, planning and implementation, measuring performance, and audit. They calculated an overall mean score for the total sample and a mean score for each of the six dimensions for each sector. Their benchmarking approach involved a comparison of each sector, using the mean score on each of the 6 health and safety management principles, with the overall sample mean across all sectors. In addition, each sector was compared to the highest performing organisation in the overall sample.

In addition to the use of mean scores and 'best in class' approaches, quartiles and/or percentiles are also another commonly used benchmarking approach. A prototypical example of this approach is HSE's Stress Management Standards. The Management Standards capture 7 dimensions of

psychosocial working conditions, which include: demands, control, manager support, peer support, relationships, role and change. Organisations that complete the Management Standards can use either of two benchmarks to compare their performance. The first benchmark is based on a 2004 survey on the psychosocial working conditions in Britain, drawing on a random sample of 1,800 people. The second, most up to date benchmark was developed based on survey data from 136 organisations. In both cases, an average score is calculated for each factor and statement contained in the Management standards. Thus, an organisation's performance is compared with either an aggregate mean score of individual responses (2004 benchmark) or a set of organisational averages for each factor and/or statement (Mackay et al., 2004).

The stress management standards approach uses percentiles to identify poor and high performers. Results of the comparison between an organisation's survey responses and the benchmark are presented using a traffic light system. In particular, the results of companies whose average on a particular factor/statement falls below the 20th percentile are presented in red (denoting that 'urgent action is required'); responses that fall on or above the 20th percentile but below the 50th percentile are presented in amber colour (denoting that there is a 'clear need for improvement'); responses that fall at or above the 50th percentile but below the 80th percentile are presented in aqua colour (denoting that an organisation's performance is 'good but there is a need for improvement'); and finally responses that fall at or above the 80th percentile are presented in green colour (denoting that an organisation is 'doing very well and needs to maintain performance').

3.2 TRANSPOSING HSCST DATA

HSL have the original benchmarking dataset generated from users of the original HSCST, consisting of over 40,000 responses. However, only limited information on the companies included in the dataset is available (e.g. company sectors).

The revision of the original HSCST consisted of removing items that had demonstrated (through statistical analyses) less discrimination, or lower loading on factor scores. This left 37 items that loaded onto eight factors. An additional 3 items were added and piloted to enhance the strength of two factors (see Sugden et al., 2009 for further information). These changes did not affect the predictive ability of the SCT, and theoretically, there was no reason why the data generated by the 37 items retained in the HSL SCT could not be used for benchmarking.

HSL successfully used this method to retrospectively analyse an organisation's trends in safety climate. Data collected by a large manufacturing organisation with a major distribution centre in 2003 and 2007 using the HSCST was transposed onto the revised factor structure, and analysed alongside data collected in 2009 using the SCT. Figure 3 illustrates the differences in factor scores across the three time points. This analysis was validated through interviews and focus groups with the organisation, which verified the trends observed. It is worth observing that the

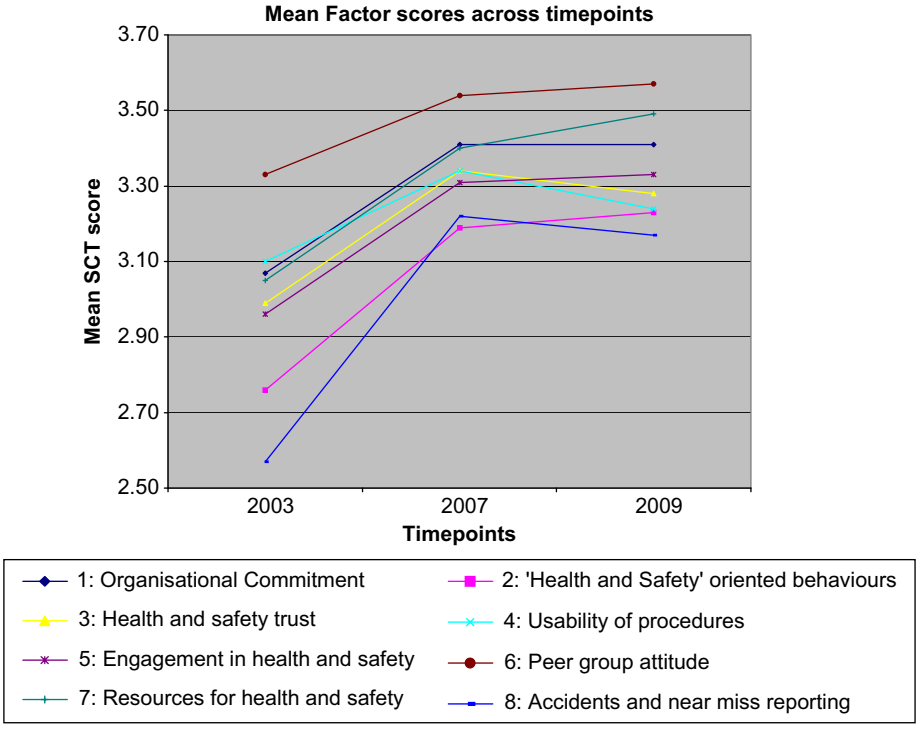

Figure 3. Analysis of 2003, 2007 and 2009 data

organisation's performance improved over the time periods, and then plateaued.

However, there are a number of limitations associated with using data generated from the HSCST for benchmarking. Despite a healthy sample of responses ($n = 40{,}000$), and information on the sectors represented in the sample, other information on the organisations that made up the sample (e.g. size of organisation) was not available. In addition, data had been generated up until 2007 (when the HSCST was removed from the market), and therefore the relevance of this data for benchmarking could be questioned. There was also a perception from users of the tool that the revised tool required 'up to date' data for benchmarking.

HSL were unable to validate the consistency of responses between the HSCST and the SCT, and were unable to verify the data (i.e. there was no way of knowing if the dataset contained replicated data). This was highlighted in feedback from users of the previous benchmarking service, who stated it was not clear whom they were benchmarking themselves against. As a result, the previous benchmarking data was no longer usable in its current form.

3.3 INDIVIDUAL BENCHMARKS VS ORGANISATIONAL BENCHMARKS

Throughout the piloting and revision of the SCT, a wealth of data was collected. Benchmarking data was calculated using individual responses (i.e. mean factor scores calculated from the total sample of individual SCT responses, across all companies) for comparison with an organisations SCT score. This was completed primarily using descriptive statistics. Analysis of these benchmarks showed the data reflected normal distributions, bar a few individual responses that were skewed. Therefore, when organisational mean scores were calculated and compared with the benchmarks, it was not possible for any of the organisational mean scores to equal the highest (or lowest) benchmarks.

The purpose of a HSL SCT benchmarking capability is to enable benchmarking analyses between working populations i.e. the focus is not on identifying individual respondents with significantly higher or lower SCT factor scores (and also, as mentioned previously, to allow organisations to compare how they perform against others, which is not possible if the focus is on individual scores). In the context of SCT scores, a group of workers who perceive they are exposed to a risk (e.g. poor usability of procedures) may generate a greater effect on SCT scores than a small number of workers who perceive they are exposed to a more prominent risk.

This also means that where large populations hold a perception (as reflected in SCT responses), even a relatively small change in the workplace (e.g. reviewing and improving existing practices) may result in more positive worker perceptions, as reflected in substantial improvements to SCT scores.

Analysis of the data suggests organisational level benchmarks (i.e. mean scores for each factor calculated across individual respondents within a surveyed organisation) are more consistent with, and better suited to the analysis of SCT scores (i.e. a representation of the general perceptions of a working population). This is consistent with the concept of 'climate' that reflects the idea of shared perceptions about organisational practices, thus requiring the aggregation of individuals responses. This approach is also similar to that used by other tools e.g. the HSE Management Standards (Mackay et al., 2004).

HSL calculated organisational SCT factor scores using a dataset of 49,204 responses (made up of the original HSCST benchmarking data and an additional dataset obtained from a construction company who had previously used the HSCST). Researchers explored the use of percentile thresholds to distinguish between performance levels (i.e. to identify whether an organisation has a good or poor result). Adjusted percentiles were used to reflect normal distributions in responses, similar to the HSE Management Standards i.e. instead of using quartiles. For each factor, percentiles were calculated based on mean factor scores, and HSL found that using the 35th, 65th and 90th percentiles generated set points, which were not clustered around the mean. For example:

- Excellent $\geq$ 90th percentile (green)
- Good $<$90th to $\geq$ 65th percentile (white)
- Average $<$65th to $\geq$ 35th percentile (yellow)
- Poor $<$35th percentile (red).

Benchmarking performance against 'good' and 'poor' performers raises awareness within organisations regarding the highest (and lowest) SCT scores currently being achieved. The most recent analyses of the HSL benchmarking dataset suggest more data is required to confirm if the use of percentiles would be a robust addition to the HSL benchmarking capability by considering the reliability, validity and generalisability of the data.

In the interim, HSL have enabled companies to benchmark their performance in a number of formats. The current highest and lowest (anonymous) SCT factor scores in the 'all industry' dataset[1] were considered a good indication of an organisations SCT performance. Presentation of this information in a spider diagram (see Figure 4) has the advantage of being clear and easily interpreted by a wide audience. This format also allows for a comprehensive comparison of an organisation's performance on all eight factors.

The aim of this ongoing work has been to explore the development of robust benchmarks. This requires consideration of the question 'how robust do benchmarks *need* to be to fulfil their purpose?'. If the aims are as previously stated (i.e. to enable comparison of performance to identify organisations achieving a positive safety culture, and facilitate improvements and adoption of practices leading to a positive safety culture) then presentation of the highest and lowest scores in the form of a spider diagram fulfils these requirements.

[1]The HSL 'all industry' dataset is a live dataset, consisting of over 30,000 responses as at March 2012.

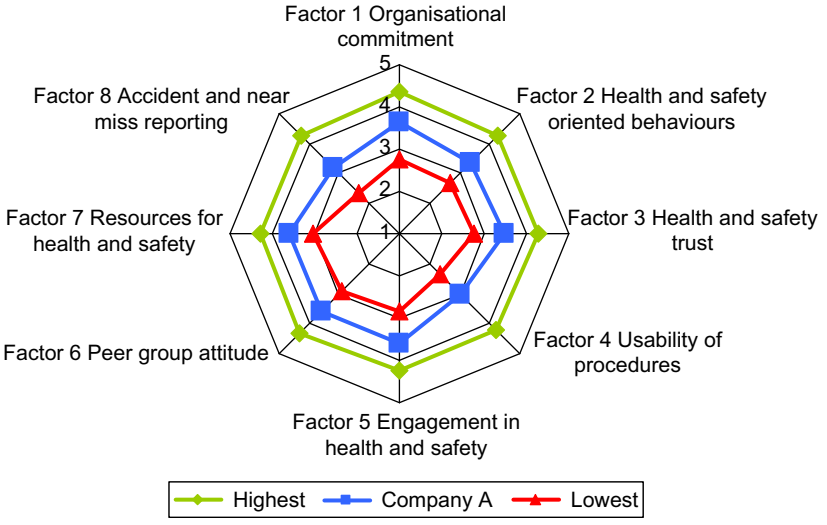

Figure 4. Benchmarking spider diagram

3.4 EFFICACY OF INDUSTRY SPECIFIC BENCHMARKS

A key user requirement identified during market research is for industry or sector specific benchmarks. Following recent work that HSL undertook with the Olympic Delivery Authority (ODA), a resounding finding was that the examples of good practice reflected the principles of good safety management (Healey/HSE in press). None of the examples identified were exclusively relevant to the construction sector, and a key implication was the potential value in transferring these practices to other industries.

In addition, the concept of 'safety culture' is not industry specific, and the distinction between 'within' or 'across' industry comparisons can be difficult. For example, large organisations may be classed under one sector, but have common working environments with other sectors (e.g. offices or production lines).

A suitable compromise may be that organisations compare 'achievable SCT scores' within sectors, but are able to share non-specific good practices across sectors for greater learning opportunities. The feasibility of such comparisons requires further exploration as the HSL dataset continues to expand.

3.5 HOW MUCH DATA IS REQUIRED FOR BENCHMARKING?

HSL's current 'all industry' benchmarking dataset consists of over 30,000 responses from 42 organisations in the UK, across a range of sectors, sizes and sites. The current dataset consists of data that has been provided by companies voluntarily, and is therefore prone to self-selection effects. Therefore, benchmarks drawn from this sample may be biased as companies that volunteer to provide their data may be more invested in health and safety, and thus are more likely to have positive safety cultures. Furthermore purchasers of the SCT may of themselves tend towards good performance. But how important is this, if the aim of the benchmarking process is for organisations to identify good practice? The dataset should be interpreted with caution, as the highest/lowest scores may not be an accurate representation of safety climate in typical companies. The assumption is that better performers will be more likely to contribute to the current benchmarking study, and therefore both the higher, lower and mean scores may be artificially inflated.

3.6 BENCHMARKING AS A PROCESS

There are a number of practical considerations associated with providing a benchmarking service. Despite limited information on benchmarking methods, there is a distinction between 'live' and 'fixed' benchmarking datasets. Benchmarking using 'live' datasets introduce the possibility of skewed data, and raise considerations over practicalities of validating data (i.e. to identify and remove replication). However a 'fixed' benchmarking dataset may not be sufficiently responsive, or may be perceived to be out of date.

Benchmarking clubs provide a forum for sharing good practice, as well as a potential source of ideas, information, methods and sharing of good practice that other organisations may benefit from (i.e. leading to more positive SCT scores being achieved).

HSL recently worked with a number of companies in the construction industry where SCT data indicated particularly positive safety cultures. HSL used interviews and focus groups to explore the practices implemented by the companies, and compiled a series of eight (SCT-factor based) case studies with the intention of sharing good practice with a wider audience (see http://learninglegacy.london2012.com/publications/safety-culture-on-the-olympic-park.php for the good practice case studies).

4. CONCLUSIONS

Benchmarking is a crucial performance improvement tool, enabling identification of best practice. Involvement in activities such as benchmarking (i.e. striving to be the best or amongst the best) is consistent with a positive safety culture.

The challenges of benchmarking can be summarised as:

- Clarifying the purpose of benchmarking:
  - Fit for purpose benchmarks – what is the overall aim of benchmarking in this context?
  - What is the most appropriate method for identifying high scores and good practice?
  - What are the benefits of trending performance over time?
- Understanding and addressing the limitations of the data:
  - How much data is required to consider the reliability of the SCT as a benchmarking measure?

Implications for data collection:

- What are the implications of benchmarking using a self-selected sample?

Further work is required to analyse the current benchmarking dataset, and to establish the efficacy of offering industry-specific benchmarks, and benchmarks based size of organisation. This will include exploring how SCT scores can map onto various safety culture maturity models e.g. Bradley curve (as cited in Krzywicki, 2011).

ACKNOWLEDGEMENTS

HSL would like to acknowledge the help and participation of the companies who have contributed to ongoing work to explore the feasibility of a HSL SCT benchmarking capability.

FUNDING

This work is supported by the Investment Research Programme of the Health and Safety Laboratory.

REFERENCES

Health and Safety Executive, 2004, Management standards for tackling work-related stress: HSE Management Standards analysis tool user manual. http://www.hse.gov.uk/stress/standards/pdfs/analysistoolmanual.pdf.

Heideggel, T., Husemann, Y., Nuebling, M., Morf, D., Siebel, T., Huth, A., Germann, R., Innerhofer, P., Faserl, A., Schubert, C., Geibinger, C., Fluckinger, K., Coi, T., and Kreienbuhl, G., 2002, Patient satisfaction with anaesthesia care: Development of a psychometric questionnaire and benchmarking among six hospitals in Switzerland and Austria. *British Journal of Anaesthesia*, 89, 6: 863–872.

HSC, 1993, ACSNI study group on human factors. *Third report Organising for Safety*. HSE Books.

HSE, 1991, Successful health and safety management. HMSO.

IAEA, 1986, Summary Report on the Post-Accident Review Meeting on the Chernobyl Accident. *International Safety Advisory Group*, Safety Series 75-INSAG-1 (Vienna: IAEA).

Keil Centre, 2002, Evaluating the effectiveness of the Health and Safety Executive's Health and Safety Climate Survey Tool. Research Report 042. HSE Books: Sudbury.

Krzywicki, R., 2011, Better safety performance through benchmarking. *Sustainable business forum*. Available at http://sustainablebusinessforum.com/bob-krzywicki/53999/better-safety-performance-through-benchmarking

Mackay, C., Cousins, R., Kelly, P., Lee, S. & McCaig, R., 2004, 'Management Standards' and work-related stress in the UK: Policy background and science. *Work & Stress*, 18, 2: 91–112.

Mearns, K., Whitaker, S. M., and Flin, R., 2001, Benchmarking safety climate in hazardous environments: A longitudinal, interorganizational approach. *Risk Analysis*, 21, 4: 771–786.

Miller, I., and Cox, S., 1997, Benchmarking for loss control. *Journal of the Institute of Occupational Health and Safety*, 1, 2: 39–47.

Rousseau, D.M., 1988, The construction of climate in organisational research. In *International Review of Industrial and Organisational Psychology,* Eds. C.L. Cooper and I. Robertson. John Wiley and Sons, Chichester.

Schein, EH, 1990, Organisational Culture. *American Psychologist*, 45: 109–119.

Sexton, J. B., Helmreich, R. L., Neilands, T. B., Rowan, K., Vella, K., Boyden, J., Roberts, P. R., and Thomas, E. J., 2006, The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. *BMC Health Services Research*, 44, 6. Available at http://www.biomedcentral.com/1472–6963/6/44

Stapenhurst, T., 2009, The benchmarking book: A how-to-guide to best practice for managers and practitioners. Oxford: Elsevier.

Sugden C, Marshall M, Binch S and Bottomley D., 2009, The development of HSL's Safety Climate Tool – A revision of the Health and Safety Climate Survey Tool. In *Contemporaty Ergonomics*, 2009. Ed. P. Bust, Taylor & Francis: London. 245–252.

Tagiuri, R and Litwin, G., 1968, Organisational Climate: Exploration of a concept. Harvard Business School, Boston.

Zohar, D., 1980, Safety climate in industrial organisations: theoretical and applied implications. *Journal of Applied Psychology* 65: 96–102.