

Quantifying changing understanding: Level 3 evaluation of leadership development interventions

The need for effective evaluation

Effective evaluation provides a vehicle for ensuring that organisations deliver the training that enables managers to grow and gain mastery in their various roles. However, evaluating the relevance of training for developing generic skills such as leadership presents an enduring difficulty for HR and training specialists. This is because it is relatively easy to obtain feedback on the participant's reaction to training, but more difficult to determine whether they have improved their understanding of leadership, or whether there is any subsequent benefit to the organisation as a whole.

Trainers might typically evaluate their courses by eliciting participant responses to course content and processes, and perhaps by examining participant retention of specific skills and strategies. These approaches are examples of Level 1 and 2 evaluation (Kirkpatrick 1998). Measuring individual development (Level 3) and organisational change (Level 4) are also necessary to gain a more complete measure of the effectiveness of a course, but they are more difficult. A number of disparate, intervening variables not associated with the course may account for an increase in participants' performance, and isolating these variables presents a challenge. This is particularly true when allowing for the impact of experience in developing leaders.

The overall purpose of our research was to measure the extent to which a participant's 'mental model' of leadership changes as a result of leadership development interventions such as education, training and experience. Argyris and Schön (1974) define a mental model as the individual's personal 'theory in use' about the world that underpins their behaviour. They may also be called 'cognitive maps', 'schemas' or 'mental constructs'.

The mental model determines the selection, interpretation, simplification and integration of information from the surrounding world. The individual relies on the mental model to make sense of the world and to formulate plans for influencing outcomes. Through experiences, a manager develops a personal mental model of leadership and it is this cognitive understanding that determines the observable leadership behaviour. We propose that a training course needs to enhance the leader's mental model of leadership if it is to improve his or her leadership performance. The purpose of this paper is to describe an aspect of our research that entails an innovative method for quantifying Level 3 changes in individuals undergoing leadership development.

Eric Stevenson

James Warn



Eric Stevenson is Assistant Director, Strategic Education and Training in the Australian Department of Defence.¹

James Warn lectures at UNSW@ADFA in the fields of Leadership and Organisational Behaviour.²

DEFINITIONS

Espoused theory: *'By 'espoused theory' we mean the theory of action which is advanced to explain or justify a given pattern of activity.*

Theory-in-use: *By 'theory-in-use' we mean the theory of action which is implicit in the performance of that pattern of activity. A theory-in-use is not a 'given'. It must be constructed from observations of the pattern of action in question.*

(Argyris & Schön 1996, p. 13)

Measuring the participant's change in behaviour

The primary purpose of an evaluation is to measure whether specific organisational and individual learning objectives have been achieved. It is generally accepted that there are four types of evaluation:

KIRKPATRICK'S FOUR LEVELS OF EVALUATION

Level 1 Reaction

Elicit participant responses to course content and processes.

Level 2 Learning retention

Examine participant retention of specific skills and strategies.

Level 3 Individual change in behaviour

Measure individual development.

Level 4 Subsequent organisational impacts

Link training to organisational change.

The Reaction Level involves measurement of participant reactions to individual sessions and to the courses as a whole; the Learning Retention Level involves a measure of participant learning; while the Individual Change in Behaviour Level often relies on 360-degree feedback to detect the extent to which supervisors, colleagues and subordinates notice any changes in an individual manager's behaviour. The fourth level, Subsequent Organisational Impacts, seeks to gain both qualitative (views of senior managers) and quantitative (numerical performance) data as a means of determining the impact of each of the courses on organisational performance.

Current methods for measuring the change in understanding are problematic. A follow-up test evaluates the effort the participant has put into reading course handouts and memorising key ideas, but does not indicate whether there has been a deeper change in understanding. Also, a number of studies have shown that knowledge retention is not a reliable predictor of longer term performance. Tests and quizzes fail to indicate whether there has been an improvement in tacit knowledge, and it is this type of knowledge that is particularly important in

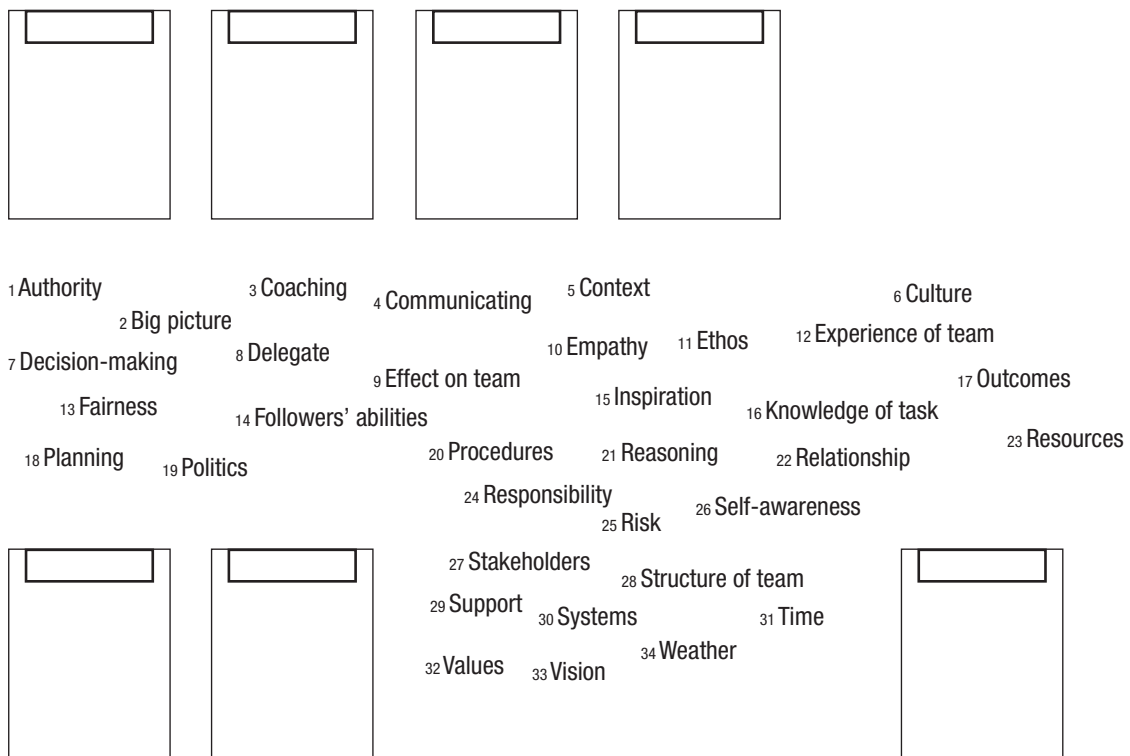
leadership. Problems also exist with 360° feedback tools. Assuming that before-and-after surveys are conducted, one might believe that it is fair to conclude that any change in behaviour can be attributed to the impact of the training. However, this claim is contentious. Perceived performance in the workplace can be influenced by a number of factors, notably motivation, resource availability and opportunity. Also, in self-ratings, the individual can learn the ideal response pattern as a consequence of being exposed to the concepts on a training course. Although 360° feedback can be a valuable tool for monitoring performance in the workplace, it may not represent an ideal tool for measuring the extent to which a course has changed understanding.

The method explored in our current research is to measure the change directly by tapping into the mental model of the leader. By constructing a picture of the manager's mental model, it is possible to identify changes in potential behaviour that are due to the individual's understanding of the world, instead of tracking observed behaviour, which is dependent on environmental contingencies.

From a cognitive perspective, leadership can be defined as the ability of an individual to effectively influence others in a given situation, based on a combination of that individual's knowledge, skills and attitudes. The effectiveness of leadership behaviour depends on an individual making correct assessments of the context, the nature of the outcomes flowing from the actions, and the needs and expectations of those involved with the event. Day (2001, p. 582) defines leadership development as, '... expanding the collective capacity of organisational members to engage effectively in leadership roles and processes.'

For leadership development courses to be considered relevant, one needs to find a high correlation between the intervention and a participant's subsequent ability to be effective in leadership roles and processes in his or her work environment. We predict that leadership development courses are likely to be more effective if they create learning environments that allow participants to develop appropriate, well-structured mental models, rather than attempting to inculcate specific leadership knowledge, skills or attitudes. Mental models allow leaders to learn complex behaviours, organise existing information and learn

FIGURE 1: STEVENSON LEADERSHIP CONSIDERATION DIAGNOSTIC(©)



new information, because they provide a framework for processing abstract concepts and skills and extrapolating them to new activities. The implication for developing leaders and managers is that many activities may not be maximising learning opportunities because of a mismatch between participants' needs and course delivery techniques.

Research participants

The participants were 907 Australian defence organisation personnel and foreign military officers undertaking Defence-sponsored courses of between one and three years duration.

The participants can be divided into three levels of leadership experience:

- **Trainee Officer level** – participants at this level were officer cadets and midshipmen undertaking a three-year undergraduate degree combined with officer training emphasising leadership development.
- **Middle Officer level** – participants undertake a 12-month course for major (equivalents) to provide them with the knowledge and skills needed in mid-career for command and staff appointments in single-service, joint and integrated environments at the operational level.
- **Senior Officer level** – participants undertake a 12-month course to provide the knowledge and decision-making skills required by senior leaders and managers to operate at the strategic level in a complex and modern security environment.

The average age for each level is Trainee (24 years), Middle (35 years) and Senior (46 years).

Research method

Our Level 3 testing involved the use of two, 20-minute activities – a questionnaire and a 'card sort'. The self-rating questionnaire contained the Bass & Avolio Multifactor Leadership Questionnaire (Avolio, Bass & Jung 1999); questions relating to participants' effectiveness as a leader in the workplace, overall satisfaction with the organisation, the learning environment on their course, and ways in which they have learnt leadership. Examples of questions on learning to lead are: I learnt to lead 'by observing others complete an activity', and '... attempting an observed activity myself'.

In the card sort, participants answer the following question:

If a friend asked you what you think about when making a leadership decision, what would you tell him or her?

A single PowerPoint slide is emailed to participants containing a diagram. In the centre are 34 words representing what people think about, around the outside are rectangles that look like clipboards (Figure 1). Participants are then asked to place each word on to one of the clipboards to make a group. There are no right or wrong answers, and participants can use as many of the clipboards as they like. They can add more clipboards, or their own words. If they do not use a word, they leave it in the middle of the page. Participants then label

FIGURE 2: TYPICAL GRAPH FOR EFFECTIVE LEADERS AT THE TRAINEE LEVEL

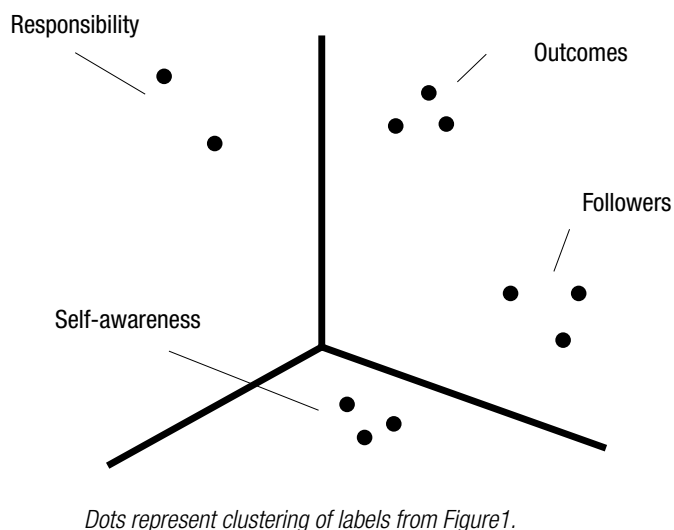
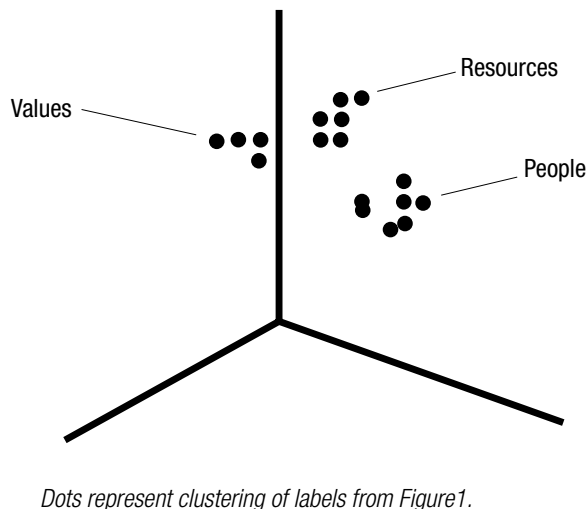


FIGURE 3: TYPICAL GRAPH FOR EFFECTIVE LEADERS AT THE MIDDLE LEVEL



each clipboard with a word that best describes the group and send it back.

The words were identified during field testing and focus groups over a number of years. The selection of the words was based on the frequency and relative importance of words used in everyday language by effective and less effective leaders at various levels, both in the military and in the broader civilian sector.

The purpose of the questionnaire was to obtain quantitative measures by which leaders could be categorised as effective or less effective, and as transformational or transactional. The card sort provides the raw data for analysing the mental model. We hypothesised that there would be systematic differences between leaders displaying different levels or styles of performance. We expected the mental models of the effective leaders to differ from those of the less effective leaders. Similarly, we expected the mental models of the transformational leaders to differ from those of the transactional leaders.

Findings of the research

The data contained in the clipboards were analysed using the multivariate techniques of hierarchical clustering and multi-dimensional scaling. Hierarchical clustering examines the responses provided by participants and, based on how similar they are to each other, divides them into common clusters. Multi-dimensional scaling enables each of the clusters of words to be plotted on a three-dimensional graph and allows visual comparisons to be made between the responses of the different types of leaders.

The three-dimensional graphs allow us to compare the relative importance of groups of words for effective and less effective participants undergoing different training and with different

levels of experience. The graphs give us an indication of what participants at a particular stage of leadership development think is important; primarily by the unique way in which they group their ideas. We found that the word clusters for leaders at a particular level tend to be closer together on the three-dimensional graph than the clusters for less effective leaders. Also, the words used to make up the clusters differ between effective and less effective leaders.

Word groups in military officers' mental models reflect a consideration of task, follower and environmental factors. The distinction becomes clearer as the level of experience increases. Both effective and less effective leaders consider the factors, although the degree of sophistication (quantity of words, relationship between clusters, logic of the grouping) increases with level. Vision, values and big picture frequently cluster in effective leaders, more so at the upper levels. The more effective leaders are able to evaluate contingency factors and to keep in mind the bigger picture.

Differences were also found between trainees after one versus three years of experience at the military training academy. In Year 1 at the Trainee level, participants tend to select all the words and map them. This may be because participants are unsure of the importance of words and it is easier to include all of them. Participants admitted during interviews that leadership is a complex process. In contrast, by Year 3, trainees have been taught a functional leadership model which serves as an heuristic for organising their evaluation of required leadership behaviour. Additionally, at the beginning of Year 3, participants undertake a major experiential leadership activity in which they are able to practise this mental model. Application of this heuristic results in simplification and a reduction in the number of words required in the

card sorts, and accounts for the Year 3 maps being qualitatively different to the Year 1 or 2 maps.

An examination of the maps of participants at each level also supports the proposition that there is a significant correlation between effective military officers and well-structured mental models. As participants progress from Year 1 at the Trainee level, through the Middle to the Senior level, the distance between and within the clusters reduces.

Effective leaders at each level tend to cluster words along strategic lines, for example Big Picture-Systems, Values-Vision; while less effective leaders tend to cluster along highly salient task lines. There is some evidence of effective leaders clustering along transformational lines, and less effective leaders clustering along transactional lines, although this evidence is inconclusive.

At the Middle and Senior level, participants scoring high transformational and high transactional grouped their words along transformational lines. This observation was also true for participants scoring high transformational and low transactional. For example, transactional participants grouped words such as coaching (normally associated with followers) with structure (normally associated with outcomes). At the Middle and Senior levels, 'values', 'vision' and 'big picture' were consistently clustered.

Participants commented after the testing was complete that they found it more difficult to 'guess' correct responses for the card sort, than in the questionnaire. This has implications for the comparative reliability of the card sort compared to a questionnaire. A degree of social desirability is inherent in the questionnaire that is not evident in the card sort. The challenge in eliciting returns on the card sort is that participants believe the technique is too simple to be able to accurately identify their mental models. As such, although the percentage of questionnaires returned at each level was similar, the more senior the participant, the less likely a completed card sort would be returned.

Although 75% of participants labelled their groups in the card sort, there was no consistency in the labels selected. This suggests that although participants can cluster words in a similar way, they may not have a shared understanding of the labels for these clusters. A failure to find common agreement on labels supports our view that the card sort is measuring tacit knowledge on leadership rather than memorised definitions.

Through learning, effective leaders develop the capacity to identify, structure and retrieve the information required for deciding on the most appropriate leadership behaviour in a particular situation. Effective leaders also have a clear, highly individualised picture of the context in which they are operating and an understanding of the implications of their actions in influencing follower behaviour. They may be able to draw on more mental 'images' more quickly, because they have a broader range of experiences. The ability to learn from experience and self-monitor the knowledge

and skills being developed assists participants in the transition from less effective to more effective leader.

We believe that the graphs can be used to:

- measure the extent to which training changes the way participants understand leadership (Level 3)
- confirm which aspects of leadership participants believe are important at their level in the workplace.

Conclusion

Evaluation is designed to ensure that the content and processes used to deliver courses meets both the organisation's and the individual participant's needs – providing learning in the most efficient and effective manner possible. However, ensuring that evaluation covers the range of learning outcomes desired is often problematic.

The emphasis on Level 3 evaluation directs attention to measuring the change that has occurred in the individual participant as a result of undergoing training or a particular leadership experience. This information is necessary in order to estimate the cost-benefit and return on investment of training in the workplace. Measuring this change directly by obtaining a picture of the participant's mental model provides a clearer understanding of the changes that have occurred for the individual, and enables other factors influencing performance to be isolated. More accurate assessments of training outcomes can only contribute to promote more effective and efficient learning and growth of leaders.

Notes

- 1 Eric Stevenson spent 24 years in the Australian Army leading various types of armoured units and military organisations, and facilitating officer education and training. Eric holds a Bachelor of Arts in Military Studies, a Master of Arts in Human Resource Development and is currently completing a Doctorate of Philosophy (part-time) with UNSW@ADFA.
- 2 Prior to his academic career Dr James Warn served as a psychology officer in the Australian Army. He is a Member of the Australian Psychological Society and a Fellow of the Australian Institute of Management. He is an accredited facilitator in leadership development tools and has consulted on organisational change programs.

References

- Argyris, C & Schön, DA 1974, *Theory in practice: increasing professional effectiveness*, Jossey-Bass, San Francisco.
- Argyris, C & Schön, DA 1996, *Organisational learning II: theory method and practice*, Addison-Wesley, Reading, Massachusetts.
- Avolio, BJ, Bass, BM & Jung, DI 1999, 'Re-examining the components of transformational and transactional leadership using the Multifactor Leadership Questionnaire', *Journal of Occupational and Organizational Psychology*, vol. 72, pp. 441-462.
- Day, DV 2001, 'Leadership development: a review in context', *Leadership Quarterly*, vol. 11, no. 4, pp. 581-613.
- Kirkpatrick, D 1998, *Evaluating training programs: the four levels*, Berrett-Koehler, San Francisco.