Exploring the Intrinsic Dimensionality of Survey Responses

I. Ali, V. Ivancevic, I. Macleod and <u>Y. Yue</u>

Behaviour and Control STC
Decision Sciences Branch
Joint and Operations Analysis Division
Defence Science and Technology Group
Email: yi.yue@dsto.defence.gov.au

Abstract: We recently completed an organisational study of a branch within a busy operational military headquarters. This paper explores methodological issues of survey design in the light of statistical analyses carried out on responses to the survey.

While at the outset of the study senior branch management perceived that various issues impacted on performance, there was no clear problem definition. In designing the study we thus needed to examine the branch's workflows, processes and interactions at the individual, team, section and whole branch level; acknowledging the importance of teamwork, this became a particular focus. With no clear statement of the problem, the survey also had to be exploratory in nature, attempting to uncover various issues that may have an effect on performance. The survey comprised seven initial themes – resources, empowerment, cognitions, culture, communication climate, information flows and teamwork – plus two further themes (strategy/goals/objectives and sections) that emerged as salient issues following preliminary analysis. In total the survey consisted of 76 statements, each of which was associated with a particular theme; the 74 respondents were asked to indicate their strength of agreement with each statement on a five-point 'Likert' scale (plus Not Applicable).

Completing a lengthy survey was an added burden for our time-poor respondents. We noticed that respondents appeared to be giving less thought to their answers towards the end of the survey and we thus endeavoured to improve our instrument design for future surveys by exploring whether similar insights might be obtained with fewer questions. As a first step we investigated latent factors underlying the survey responses, finding that three dominant orthogonal factors accounted for 83.5% of the total variance and that each of these factors was heavily loaded on a single theme. The intrinsic dimensionality was thus much smaller than our choice of nine themes suggested. We further found that respondents' average choices for the three themes associated with the dominant factors could be used to estimate (better than chance) their averages for the remaining six themes. The variance of the actual averages here was 0.49 (on the Likert scale of 1 through 5) versus 0.21 for the variance of the difference between the actual and estimated averages.

There were unexpectedly large pair-wise cross-correlations between some of the themes. Part of this could be attributed to the interdependencies between themes revealed by factor analysis. A further contribution apparently arose from the tendency of some respondents to 'cluster' their answers within a rather limited range of the Likert scale in the latter part of the survey (as if they were giving less thought to their answers). A non-parametric two-sample Kolmogorov-Smirnov test showed that the range of answers for the first half of the survey was significantly greater than the range for the second half ($p < 1\%$). This tendency towards greater uniformity of answers as respondents progressed through a lengthy survey was previously observed by Galesic and Bosnjak (2009), who further noted that participation rates tend to decrease as surveys lengthen.

Our explorations suggest that with complex surveys, and when the research questions are not already well-defined, the dimensionality of the situation being investigated may well be somewhat smaller than implicitly assumed. In this case, if through pilot studies the most-important and relatively independent themes can be identified, then the number of questions may be able to be reduced without compromising the survey outcomes. Indeed, with time-poor respondents there may be a trade-off between the quality of responses and the length of the survey. Out of the 76 statements in the current survey, for example, only 30 of these fell within the themes on which the three dominant factors were most-heavily loaded.

Keywords: *Survey design, survey analysis, organisational study, factor analysis*

1. INTRODUCTION

Recently we completed an organisational study of a branch within a busy operational military headquarters. The overall study included a diagnostic phase, to identify the most important issues that were affecting performance, and an intervention phase, to evaluate and implement potential remediation measures. The diagnostic phase consisted of surveys as well as modelling and simulation. This paper relates to one of the survey instruments in the diagnostic phase and explores methodological issues of survey design in the light of the results of statistical analyses carried out on the survey responses.

At the outset of the study senior branch management perceived that work imbalance between different teams/organisational sections, process inefficiencies, lack of clear objectives, sub-optimal management of crises, etc. impacted on performance of the branch, but there was no clear problem definition. Therefore, while designing the branch study it was essential to cast a wide net and examine the branch's workflows, processes and interactions at the individual, team, section and branch level. The operational tempo meant that branch staff had heavy workloads and we had to balance the survey length against the level of detail sought.

2. DESIGN OF THE SURVEY

At the productive core of the branch is its various teams; other staff allocate work, monitor, guide and support these teams. Teamwork was thus a particular focus of the study. Teams are situated in a broader organisational context, performing multiple functions and working towards overall organisational goals. Mathieu *et al.* (2008) – in their review of a decade of team effectiveness literature – show that blended composite measures are far better indicators of overall team effectiveness compared to those that assess only one aspect of performance. Kozlowski and Ilgen (2006) identified key team processes and emergent states that influence team effectiveness. In broad categories these constructs are: team affective-motivational processes, team behavioural processes and team cognitive processes. The attitudes, behaviours and cognitions are referred to as the ABCs of teamwork where: attitudes are the affective attributes necessary for effective team performance; behaviours are the skills and procedures needed for teamwork; and cognitions are the necessary elements of knowledge and experience for teams to be effective (Salas *et al.* 2009). However, each of these constructs has different dimensions that individually or in combination contribute to team effectiveness. The different aspects of the ABC constructs shown in Table 1 served as one of the considerations guiding development of the branch survey.

Table 1. Aspects of ABC constructs underlying team performance[1].

Attitudes (Motivational-affective processes and states)	Behaviours (Behavioural processes)	Cognitions (Cognitive structures and processes)
Cohesion	Communication/Interaction for information exchange	Shared mental models
Trust	Coordination	Shared knowledge
Collective efficacy	Cooperation	Shared problem model/ Situation awareness
Collective orientation/Belief in the importance of teamwork	Problem detection and solving	Learning
Goal commitment	Assertiveness (task related)	
Empowerment	Adaptability	
Psychological safety	Leadership (shared/distributed) and initiative	
Interpersonal relations	Conflict management	
Reward attitude	Mutual motivation	
	Mission analysis	
	Back-up/supportive behaviour	

Given the lack of a clear problem definition, the survey had to be exploratory in nature, attempting to uncover various issues that may have an effect on individual performance and overall branch effectiveness. The survey comprised seven themes: *resources[2], empowerment, cognitions, culture, communication climate, information flows* and *teamwork*. Within each theme the questions probed aspects of the ABC constructs for

---

[1] There is no relationship between aspects implied by the shaded bars – these serve simply to improve readability.

[2] Where terms are italicised, they refer to themes in the specific set used herein as opposed to normal usage.

assessing team effectiveness. In total the survey consisted of 76 statements (e.g., "The current work I am doing challenges my thinking"); the respondents were asked to indicate the strength of their agreement with each statement on a five-point 'Likert' scale – Strongly Agree (numeric value = 5), Agree (4), Undecided (3), Disagree (2), Strongly Disagree (1) – plus a Not Applicable (N/A) category. Initial analysis indicated that in addition to the above themes, strategy/ goals/objectives and sections emerged as salient aspects[3]. These were added to the seven survey themes for subsequent statistical analysis. The value for each theme and each respondent was calculated by adding their numerical answers (1 through 5) and dividing by the number of statements pertaining to each theme[4], as follows:

1. *Resources* (statements 8, 9, 10, 26, 27, 28, 32, 33, 47, 55),
2. *Empowerment* (statements 5, 11, 15, 16, 29, 45),
3. *Cognitions* (statements 3, 4, 6, 58, 64),
4. *Culture* (statements 1, 2, 14, 20, 34, 37, 59, 62, 63, 66, 67, 70, 76),
5. *Communication climate* (statements 12, 13, 17, 25, 35, 50, 53),
6. *Communication plus information flows* (statements 7, 31, 56, 57, 65, 69, 71, 72),
7. *Teamwork* (statements 39, 40, 41, 42, 43, 44, 48, 49, 51, 52, 54, 60, 61, 68, 73, 75),
8. *Strategy/goals/objectives* (statements 18, 21, 36, 38) and
9. *Organisational sections* (statements 19, 22, 23, 24, 30, 46, 54, 74).

While the survey questionnaire may appear to be lengthy, this was necessitated by the exploratory nature of the study and the multiple levels of analysis.

3.   ANALYSIS OF SURVEY RESPONSES

We were mindful that completing a lengthy survey was an added burden for time-poor respondents. On the five-point scale the range of answers in our survey appeared to be widest with the early statements. We proceeded to test whether the distribution of answers in the first half of the survey was significantly different from that in the second half. We first assessed each of these distributions for Normality, using the popular Kolmogorov-Smirnov test in its one-sample form. As neither distribution was found to be consistent with Normality, we then conducted a non-parametric comparison of the two distributions using the two-sample Kolmogorov-Smirnov test. This comparison showed that the spread of answers for the first half of our survey statements was significantly greater than that of the second half at $p < 1\%$. Such an increased uniformity of answers with later questions in a survey has previously been observed by Galesic and Bosnjak (2009).

The survey yielded useful information, highlighting several areas for potential improvement. However, the apparent tendency for our respondents to give less thought to their answers in the latter part of the survey motivated us to improve our instrument design for future studies. To this end, we wished to explore whether similar insights could have been obtained with fewer questions.

3.1.   Factor Analysis of the Survey Response Space

First, to see whether the dimensionality of the situation we were studying was as large as our choice of nine themes implicitly suggested, we investigated latent factors underlying the survey responses. The three dominant orthogonal factors shown in bold text in Table 2, each heavily loaded on a single theme, accounted for 83.5% of the total variance, indicating that the intrinsic dimensionality of our response space was closer to 3-D than 9-D.

3.2.   Using the Three Dominant Factors to Estimate Responses with Other Themes

Fortuitously, each of the three dominant factors was heavily loaded on a single theme, respectively Theme 7 (*teamwork*), Theme 1 (*resources*) and Theme 8 (*strategy/goals/objectives*) for Factors 1 through 3. Confirmation of the dominance of the three factors and the extent of their alignment with three of the themes was obtained by showing that each respondent's average answers for these themes could be used to estimate (better than chance) their averages for the remaining themes (2, 3, 4, 5, 6 and 9). We take Theme 5 and Respondent 23 as an example to illustrate how an estimate can be calculated. Respondent 23's average answers to the questions aggregated in Themes 7, 1 and 8 were 2.27, 2.40 and 4.25 (termed *R1* through *R3* below), and the loadings for Factors 1 through 3 on Theme 5 were 0.53, 0.30 and 0.64 (termed *L1* through *L3*). Because we are working in an approximately orthogonal 3-D space, we use the square root of sums of squared distances on each axis to compute lengths in this space. To normalise the vector of loadings for

---

[3] On reflection, we realised that the theme of *strategy/goals/objectives* could be seen (at least in part) as an outcome of the survey rather than as a variable.

[4] We used averages here rather than medians or modes, given that with our five-point scale either of the latter measures would lead to a significant loss of information through quantisation.

Theme 5, we first compute scaling factor S as $1/\text{SQRT}(L1^2+L2^2+L3^2) = 1.13$. Our estimated average for Theme 5 is then $S*\text{SQRT}((L1*R1)^2+(L2*R2)^2+(L3*R3)^2) = 3.46$, compared with the actual average of 3.00.

Table 2. Factor Loadings on Themes.

		Factor 1	Factor 2	Factor 3
	1	.181	**.940**	.252
	2	.570	.268	.632
	3	.670	.300	.388
	4	.629	.244	.635
Theme	5	.525	.302	.636
	6	.631	.276	.548
	7	**.926**	.100	.175
	8	.196	.236	**.892**
	9	.679	.098	.615

Referring to Figure 1, with the possible exception of Theme 4 the overall correspondence between the estimated and actual averages is reasonable. Repeating this process across all 74 respondents and the remaining 6 themes, the variance of the actual averages was 0.49, and the variance of the difference between the actual and estimated averages was 0.21. The mean of the actual averages was 2.45 and the mean of the estimated averages was 2.50.
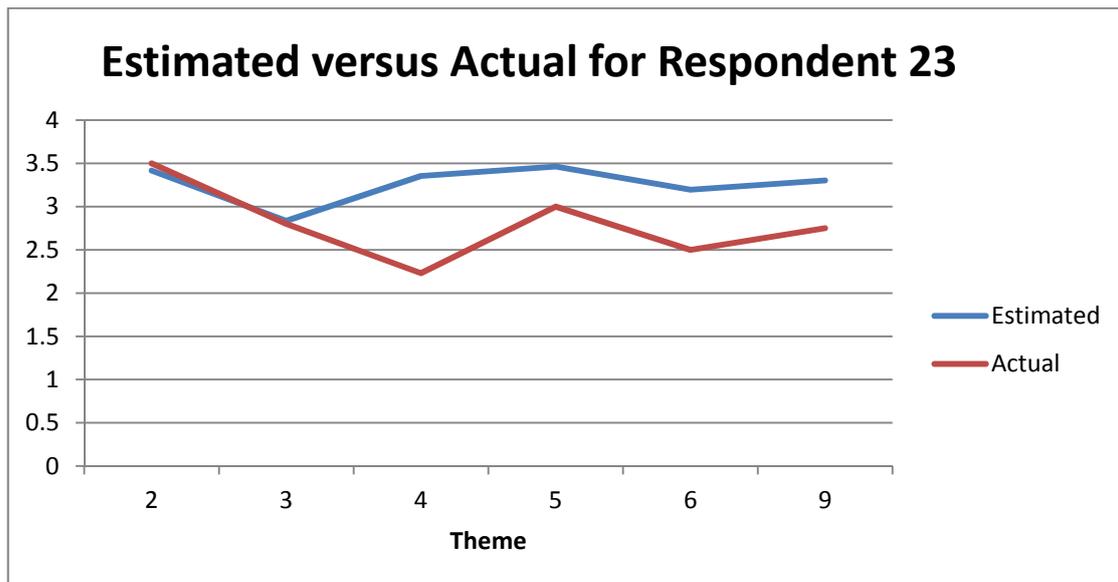

Figure 1. Estimated versus actual averages.

3.3. Pair-Wise Cross-Correlations Between Themes

As a result of interdependencies between the themes, as revealed by the factor analysis, there were significant theme to theme cross-correlations (across the 74 respondents) – see Table 3. While part of the observed cross-correlations resulted from understood real relationships between themes (discussed below), another part could well have resulted from an apparent tendency of an appreciable fraction of our respondents to 'cluster' their answers within a rather limited range of the Likert scale as they progressed through the survey. The standard deviations of these respondents' answers to the 76 survey statements were typically less than 1.0 on the Likert scale ranging from 1 to 5 (reducing to standard deviations of typically less than 0.5 when the

statement answers were aggregated into themes)[5]. For such respondents, this inclination towards increasing uniformity of answers may have resulted from the effort required to complete a lengthy survey, such that (possibly as a result of time pressure or fatigue) they gave more thought to early rather than later statements. As noted above, across all respondents the spread of answers to the first half of the survey was significantly greater than that of the second half ($p < 1\%$).

Table 3. Cross-correlations between themes.

Theme

	1	2	3	4	5	6	7	8	9
1	**1.00**	.52	.45	.50	.51	.50	.33	.49	.41
2	.52	**1.00**	.69	.81	.75	.68	.67	.71	.79
3	.45	.69	**1.00**	.69	.66	.69	.63	.59	.65
4	.50	.81	.69	**1.00**	.85	.80	.67	.67	.85
5	.51	.75	.66	.85	**1.00**	.74	.60	.66	.70
6	.50	.68	.69	.80	.74	**1.00**	.68	.68	.76
7	.33	.67	.63	.67	.60	.68	**1.00**	.43	.75
8	.49	.71	.59	.67	.66	.68	.43	**1.00**	.69
9	.41	.79	.65	.85	.70	.76	.75	.69	**1.00**

(Theme — row axis label)

Even allowing for the effect of 'clustering' of answers, the extent of cross-correlations between themes revealed in Table *3* was interesting, with the smallest cross-correlation of 0.33 between *resources* and *teamwork* and the largest values of 0.80, 0.81, 0.85 and 0.85 (again) between *culture* and (i) *communication plus information flows*, (ii) *empowerment*, (iii) *communication climate* and (iv) *organisational sections*. The pair-wise cross-correlations between the latter four themes are similar to the products of their individual cross-correlations with *culture*. For example, the cross-correlation of 0.76 between Theme 6 (*communication plus information flows*) and Theme 9 (*organisational sections*) is not very different from the product of their cross-correlations with *culture* (0.80*0.85 = 0.68), suggesting that any apparent relationships here may largely be derived from the individual relationships with *culture*.

We return to the question of real relationships between themes and, in particular, the involvement of Theme 4 (*culture*) in all four of the largest cross-correlations. The literature on organisational culture posits that culture is a unique organisational feature – an explicit product emanating from social interaction – and it develops over time in response to the organisational environment (Schein 1993). Therefore, culture can be assumed to be one of the most significant aspects of organisational life influencing the behaviour and relationships of individuals and groups within an organisation. This importance was reflected in the number of our survey questions pertaining to aspects of culture (13 out of 76, second after teamwork, another very important aspect, which had 16). Culture provides, or fails to provide, the social energy that moves people to act.

Organisations consist of subgroups that have specific characteristics together with their own sense of identity and subcultures; these may differ from the overall organisational culture. Research also points to the impact of organisational culture on the communication climate and information flows within organisations (Warne *et al.* 2003; Ali 2011). An open and supportive communication climate, or otherwise, is a natural outcome of an organisation's culture. Culture is thus an important aspect to take into account when researching organisational phenomena (Hall 2005). It needs to be investigated at a wide organisational level as well as at a subgroup level. In the current research we did so by exploring cultural aspects at the branch, organisational section and team level. In the light of the literature here, the fact that the four highest cross-correlations among the themes involved *culture*, is not in itself surprising. The magnitude of these cross-correlations was, however, greater than might be expected.

---

[5] The averages across the themes for each respondent here ranged from 1.25 to 4.4, meaning that the variation between a respondent's averages for each theme was typically much less than the variation between averages across respondents.

## 4.	DISCUSSION AND CONCLUDING REMARKS

Given that culture appears to be such a central aspect of organisational life, the question might arise as to why it wasn't one of the themes onto which a dominant factor was heavily loaded. The reason here is that *culture* was the theme most commonly correlated with other themes, whereas exploratory factor analysis seeks to find new variables that are uncorrelated with each other.

Our investigations indicated that because of interdependencies between the questionnaire themes, the space formed by answers to the survey statements was much less complex than that implied by our choice of nine themes. If we could have found a smaller set of relatively independent themes and formulated a set of statements that adequately probed them, then our survey could have been considerably shorter without compromising the outcomes. The challenge here is, of course, that such themes may not correspond to conventional ways of describing and thinking about organisational entities, making it difficult for researchers to assemble an appropriate set of statements and for survey respondents to understand them.

The success in using three themes (based on answers to only 30 out of a total of 76 statements) to estimate average answers for the other six themes is encouraging. The explorations described herein are very much a work in progress; it is not yet clear how generalisable the results might be to different organisational settings. To the extent that surveys can be reduced in length through pilot studies there may be a twofold payoff: first, the quality of the survey results may improve if respondents can think equally seriously about all statements; and second, as noted by Galesic and Bosnjak (2009), the participation rate with respect to the number of potential respondents who complete the survey may also improve.

In the interest of study reproducibility, the authors are willing to make the data gathering instrument used in this research available upon request. The data gathered in the course of this study is, however, for official use only and must remain confidential.

REFERENCES

Ali, I. (2011). Coexistence or operational necessity: the role of formally structured organisation and informal networks during deployments. *16th International Command Control Research and Technology Symposium (ICCRTS)*, Quebec City, Canada, 21–23 June 2011.

Galesic, M., and M. Bosnjak (2009). Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly*, 73(2), 349–360.

Hall, M.L. (2005). Shaping organizational culture: a practitioner's perspective. *Peak Development Consulting* II(1).

Kozlowski, S., and D. Ilgen (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest* 7(3), 77–124.

Mathieu, J., M. Maynard, T. Rapp, and L. Gilson (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management* 4(3), 410–476.

Salas E., M. Rosen, C. Burke and G. Goodwin (2009). The wisdom of collectives in organizations: an update of the teamwork competencies. In: Salas E., G. Gooodwin and C. Burke (Eds.) *Team effectiveness in complex organizations: cross-disciplinary perspectives and approaches,* pp.39–82. New York, Psychology Press, Taylor and Francis Group.

Schein, E.H. (1993). Organizational culture and leadership. *Organizational Dynamics* 22(2), 40–51.

Warne, L., I. Ali, and C. Pascoe (2003). *Social learning* and *knowledge management – a journey through the Australian Defence Organisation: the final report of the Enterprise Social Learning Architectures Task.* Canberra, ACT, Defence Science and Technology Organisation.