THE INTERNATIONAL JOURNAL OF BUSINESS & MANAGEMENT

An Alternative to the Likert Scale When Polling Employees

Jan Van De Poll
Managing Partner, Department of Tlab Research
Transparency Lab, Amsterdam, Netherlands

Abstract:
The input from employees is a crucial source of information for strategic decision-making in organizations. The Likert scale is the scale most often used for employee polls but results in a variety of biases - including socially acceptable answers – influencing the overall scores. Hence, to offer an alternative, we improved a Guttman scale specifically for employee polling (Guttman-Poll). This improvement asks employees about verifiable facts and -behavior, taps actual situation and the employees' ambition, caters to target setting, and provides additional managerial insights into, e.g., organizational alignment and knowledge sharing. We compared this scale to a Likert scale in 5 different employee polls. The answers following a Likert scale were normally distributed but were significantly more biased to the higher end of the scale: almost 80% of questions scored higher than five on a scale from zero to ten. The answers on the Guttman-Poll scale were normally distributed across the entire scale. The Guttman-Poll scale delivers relatively noise-free input (tallying verifiable facts/-behavior rather than averaging opinions), which may drive algorithms and A.I.

Keywords*: Employee polling, survey design, Likert, Guttman-Poll, alignment, knowledge sharing*

1. Introduction

Every organization takes strategic decisions that require change management: a combination of a sense of urgency, a vision to follow, the empowerment of the organization, planning for wins, and institutionalizing new approaches to make these decisions a reality (Kotter, 1995). Yet, organizations realize that in recent years strategic decision-making is getting ever more complex (Conteh, 2013) and that the time available for decision-making is getting increasingly shorter (Kotter, 2012). Polling the aggregated views from large numbers of individuals – the 'wisdom of the crowd' – has proven to outperform financial market models (Giles, 2005) as well as models in other areas, provided that the crowd meets specific requirements: diversity, independence of its members, decentralization, and some mechanism for data aggregation (Surowiecki, 2005).

If there is one regular feature in asking large groups of employees, it is the employee engagement poll. Predominantly, the question format asks for opinions using Likert scales (Langford, 2009; Trotman, Tan, &Ang, 2011; Joshi &Sodhi, 2011; Bailey, Madden, Alfes, & Fletcher (2017), e.g., 'Overall, I find the work that I do meaningful' (Mone& London, 2014). Likert scales have been designed to measure attitudes (Likert, 1932). Asking for agreement on, or asking for an applicability score of, statements may require corrections for response latitudes (as a sign of respondent non-involvement) (Lake, Withrow, Zickar, Wood, Dalal, &Bochinski, 2013), for a variety of response styles (De Jong, Steenkamp, Fox, & Baumgartner, 2008), for various sampling errors (Piterenko, 2013) and multiple types of bias (Roulston& Shelton, 2015) all hampering the unambiguity of the employees' input. Statements do not make it easy to come up with an ambition score or allow management to set a target to drive the change management either.

2. Objective

The purpose of this paper is to test an alternative scale for employee polls to obtain the feedback of large amounts of employees about their view regarding strategic options provided by the organization's management. This scale must capture verifiable facts and verifiable behavior about the organization's actual situation as well as the employee's ambition on what and how much to change. The scale must also allow setting targets. Statistically, the scale must be normally distributed around the middle of the scale. After all, a scale that yields, e.g., scores towards the higher end of the scale, might erroneously paint a too rosy picture. Finally, we want to investigate whether this alternative scale delivers additional managerial insights not possible with Likert scales.

3. Method

3.1. Procedure and Participants

We first designed an alternative scale based on the Guttman scale (Stauffer et al., 1950), specifically designed for employee polling ("Guttman-Poll"). Next, we investigated how Likert- and Guttman-Poll scales scored in comparison to each other (mean, standard deviation, kurtosis, skewness, and the percentage of employee answers scoring higher than 5 in a range from 0-10). Therefore, we analyzed several employee polls in a multinational construction company with a total of 39,507 respondents submitting 632,112 answers over the course of 26 different employee polls during 2,5 years. These

employee polls were entirely based on a Likert scale and about strategic employee matters (personal development, salary, management style, culture). Sample questions include:

- (*) "People in this organization are treated fairly."
- (*) "My organization provides me with the tools to do my job."
- (*) "I get the information I need to do my job."
- (*) "I am satisfied with my supervisor."

Next, we compared two employee polls in one division of this construction company. One poll was based on a Likert scale, the other on a Guttman-Poll scale, and again referring to (strategic) employee matters. Employees submitted 10,820 answers over the two polls combined, consisting of 8,320 responses on Likert questions and 2,500 responses on Guttman-Poll questions.

We wanted to exclude an influence on the comparison between Likert and Guttman-Poll because of differences in the audiences answering the polls. Hence, we subsequently conducted four other employee polls in other organizations where the polls were a mix of Likert- and Guttman-Poll based questions and answered by the same employee audience. With these four polls, we collected another 38,219 responses divided over 7,501 responses on Likert questions and 30,718 responses on Guttman-Poll questions. These four additional employee polls were about a line-staff assessment (business units rating the service of the staff departments) in a Dutch industrial services organization, about a measurement of the deployment of a strategy in a Dutch care organization, about an office productivity assessment in a Dutch energy company and the change management around an Enterprise Resource Planning (ERP) implementation in a Dutch academic hospital. The number of responses collected in the seven abovementioned employee polls only refers to answers about the actual situation in the organization.

3.2. Measures

For strategic decision-making in a change management setting, the wisdom of the crowd is aggregating each individual's view on how to change for upper management to decide on a course of action. That means we must capture – for each employee and each question in the employee poll – four different components: (1) an actual situation, (2) a preferred position (preferred by the respondent), (3) a management target (the situation chosen by management), and (4) a time frame. The smallest part of an organization's change management 'vector' is the individual vector of each employee's answer on a single survey question. Add all these responses together – by summarizing these individual vectors – and arrive at the aggregated ambition vector per question, per team, per management level, per job role or for the entire organization. As we can safely assume that management will not copy the employee ambition for 100% (for example, because of budget constraints or limited organizational capacity), this ambition vector should be corrected for the impact of upper management's final decision on how to proceed.Guttman scalability measures' breaking points' (Uhlaner, 2002): each next answer represents a higher quality level with which the question has been executed. These levels allow for target setting. This scale design is different from the Longitudinal Guttman Simplex (for example, Hays &Ellickson, 1991), which compares response patterns of a Guttman-designed questionnaire (the 'Now'-score) in time. Yet, a Guttman scale's answer options might refer to levels of agreement with statements too, and therefore still measure attitudes rather than verifiable facts and -behavior, e.g.:

Q. How will Project XYZ increase our sales/revenues/profits?

- Answer 1: it will NOT affect our sales/revenues/profits
- Answer 2: it will NOTABLY affect our sales/revenues/profits
- Answer 3: it will DRAMATICALLY affect our sales/revenues/profits

What we wanted to know in the employee poll is: exactly how far is an employee on a specific aspect of the strategy? What does he plan to improve in the foreseeable future? Would that ambition fit with the view of upper management? Guttman scaling works with current-status data (a term used by Diamond, McDonald, and Shah, 1986) and can also cover the Time dimension. For example, *now* the team does not have team objectives, but they will *in 6 months*. We added the time dimension in the Guttman scale in the following way to represent the vector of an employee's answer to a single question:

Q. How have you defined your team objectives?

	Now	In 6 months
1. We have no team objectives (yet)	☐	☐
2. We have a qualitative description	☐	☐
3. We have formal, SMART key performance indicators.	☐	☐

The answers given by the employees are objective and verifiable in such a way that they could classify as 'objectively real' or 'a testable proposition' (Ahrens & Chapman, 2006). This is because we abstain from adjectives and adverbs to reduce interpretation bias and insert words like, e.g., 'formally,' 'regularly,' 'periodically,' 'documented' and 'described' to reduce respondent self-report bias (a problem raised by Donaldson and Grans-Vallone, 2002) and to help with verification and to ensure objective answers and eliminate employees adding a cognitive or emotional meaning (cf. Frese& Zapf, 1988). Given the objectivity requirement for 'binary (no/yes), numerical or categorical representations' (Plewis& Mason, 2007, p. 192), we could say that this way of formulating answers is a 'no/yes' check for categories: "Yes, we have qualitative objectives but, no, these objectives are not SMART." Yet, there indeed remains subjectivity. The questions and answers are subjectively chosen. They usually represent the strategic reality as perceived by upper

management or the consultant firm that creates the questionnaire. Then, the best answer, as used in a question, is not theoretically the best answer. For example, an even better answer than Answer 3 in the above team objectives question is, "We have formal, SMART key performance indicators *that get formally reviewed every three months*." There might be a need for more than three answers in some cases. The 'In 6 months' is just a placeholder: it could be any moment in time later than 'Now.' And how objective and verifiable we try to make the actual score, the ambition score in 6 months will always be a respondent's opinion.

3.3. Data Analysis

The Likert scales in each employee poll were of a Disagree / Slightly disagree / Neutral / Slightly agree / Agree nature. For scoring purposes, these answers were rated 0, 2.5, 5, 7.5, and 10, respectively. Each Guttman-Poll question had three answers. The 'worst' answer of three (the current situation) was rated with a score of 0—the 'middle' answer (the intermediate step) with a score of 5. And the' best' answer (reflecting the content of the strategy that needed to be achieved) with a score of 10. An average score for a dimension (a group of questions covering one topic of the management's strategy) was achieved simply by averaging the scores of the individual questions. Total scores were based on particular questions, too, to avoid averaging averages. We abstained from any form of weighting among these questions and answers (Likert and Guttman-Poll alike) for three reasons. Firstly, we deemed the choice of questions and answers to include in the questionnaires a far greater error than the error that would be solved with weighting. Secondly, the effort of determining the proper weights would require a very considerable time effort that was at odds with the time available to execute these employee polls. Thirdly, management usually wants to focus the managerial attention, not on a discussion on weights but a discussion of (re-)prioritization and implementation steps. Skipped questions were not considered in the calculation of the scores (as we assumed the question's non-applicability for that individual respondent). The same scoring methodology was applied in the Guttman-Poll questions for the Ambition ('In 6 months') score.

4. Results

4.1. Comparing Scales

We first analyzed how the six employee polls scored.

	Likert						Guttman-Poll					
	N	Mean	StDev	Kurt.	Skew.	>=5	N	Mean	StDev	Kurt.	Skew.	>=5
One employee poll, using a Likert scale												
Employee matters	632,112	7.60	0.49	1.20	-0.71	100%						
Two comparable employee polls, one Likert and the other Guttman-Poll, in the same organization												
Employee matters	8,320	7.87	0.69	-0.29	-0.53	100%	2,500	5.60	1.52	0.92	-0.61	72%
Four employee polls with Likert and Guttman-Poll questions combined in one and the same poll												
Line-Staff assessment	2,309	5.88	0.95	1.22	-0.72	90%	8,848	4.21	1.26	-0.81	0.43	23%
Strategy deployment	2,587	6.51	1.04	-0.39	0.49	100%	9,985	5.35	1.07	0.12	0.29	66%
Office productivity	230	5.74	1.05	-1.15	-0.39	75%	403	5.11	1.22	0.47	1.15	40%
ERP implementation	2,375	4.74	0.73	0.06	-0.67	67%	11,482	4.04	1.66	-0.94	-0.04	42%
Group	7,501	5.78	1.08	-0.28	-0.03	78%	30,718	4.51	1.45	-0.40	-0.01	43%
% of total	20%						80%					

"N" = number of questions answered. "St.Dev." = Standard deviation. "Kurt." = Kurtosis.
"Skew." = Skewness. ">=5" = percentages of questions scoring 5.0 or higher
"% of total" = the percentage per scale of the total number of questions answered

Table 1: Comparison: How Question Scales Difer

The first employee poll in the construction company was only using Likert scales and was the poll that raised our concerns. An average score of 7,60 with a standard deviation of only 0,49 and 100% of questions scoring above 5,0!

Perhaps a top-performing organization? Longitudinally, it turned out that this average score was roughly stable during the 2,5 years of measurement despite notable changes in terms of management style, culture, and information provisioning to employees.

The second and third employee polls were conducted within one division in the same construction company. With comparable topics polled, the Likert version scored an average of 7,87 with a standard deviation of 0,69 and again 100% of questions scoring above 5,0. The Guttman-Poll version, though, had an average score of 5,60 with a much higher standard deviation of 1,52 and 72% of questions scoring above 5,0.

To exclude any influence in scores due to differences in the employee audience, we conducted four more employee polls that contained both Likert- and Guttman-Poll scales. Other than average score, standard deviation, and the percentage of questions scoring a 5,0 or higher, we analyzed both skewness and kurtosis. "A general guideline for

skewness is that if the number is greater than +1 or lower than –1, this is an indication of a substantially skewed distribution. For kurtosis, the general guideline is that if the number is greater than +1, the distribution is too peaked. Likewise, a kurtosis of less than –1 indicates a distribution that is too flat." (Hair et al., 2017, p. 61). Table 1 shows that the Likert questions scored on average higher (5, 78) than the Guttman-Poll scale (4,51) with a smaller standard deviation (1,08 versus 1,45) and a much higher percentage of the percentage of questions scoring a 5,0 or higher (78% versus 43%). Comparing skewness and kurtosis showed that both scales could be considered normally distributed. Hence, the answers in the polls based on a Likert scale were normally distributed but were significantly more biased to the higher end of the 0-10 scale. The answers on the Guttman-Poll scale were normally distributed across the entire 0-10 scale.

4.2. Additional Insights When Using Guttman-Poll

The Likert scale shows the actual situation of the employees' attitudes towards certain statements. The Guttman-Poll scale replaces attitude by verifiable facts or -behavior. It adds a time dimension to cater to employee ambition and management target, which opens up a lot of new, beneficial insights. When comparing the scores for the actual situation with the ambition scores (say, in 6 months), it is possible to understand in what areas the respondents see the organization's improvement priorities. Then, to understand organizational alignment, it becomes possible to see how the ambition of different groups compares. Does management want to go to the right and the work floor to the left? Is everyone in agreement on topic A but very much divided over topic B? When it comes to knowledge sharing, Respondent 1 needs to improve on question A but can ask Respondent 2 – who already scores the target answer - for assistance. Conversely, Respondent1 could help Respondent 2 on question B. Considering all respondents and all the questions in the employee poll, this leads to an organizational network.

<< Insert Figure 1 around here >>

5. Discussion

A Likert scale is a cornerstone of academic research, and its merits are undisputed. Yet, when focusing on its application to tap factual employee input for strategic decision-making, there are evident shortcomings—biased towards the positive scores, no objective input, and no possibility to capture a specific ambition or set a target. Our alternative Guttman-Poll scale design solves these shortcomings.

We compared employee polls with Likert- and Guttman-Poll scales in 6 employee polls in 5 different industries to assess the scales' statistical distribution. With the available employee polls at hand, we concluded that – when applied for tapping factual employee input to benefit strategic decision-making – Likert scales seem to have a so-called "positivism bias": a tendency to paint a rosy picture. Why have a scale from 0-10 if 78% of scores are higher than 5.0 (counting our four mixed Likert/Guttmann polls) or 99% of scores are higher than 5.0 (the Likert scores on all six polls combined)? The overall results seem to reflect an acquiescence response style (De Beuckelaer, Weijters, &Rutten, 2009). Consequently, if the vast majority of scores (if not all) are above 5.0, there is little incentive to improve things.

In terms of internal versus external validity, the Guttman-Poll scale scores very well in terms of the connection between our study results and a 'belief of reality' (cf. Koro-Ljungberg, 2008). Concerning external validity (could the findings be applied to other situations), we see no reason why the benefits of the Guttman-Poll scale couldn't be replicated for strategic decision-making in other organizations.

6. Limitations and Future Research

At the same time, there are several cautionary remarks to be made about our comparison and about the design of the alternative Guttman-Poll scale, which also fuels the need for more research.

Firstly, the choice of questions is – despite being well-intended – scientifically arbitrary. This arbitrary choice holds for both the scales. The same also holds for the answers in the Guttman-Poll scale. In that light, the choice of what described the 'worst' answer and what described the 'best' answer was also arbitrary. It was possible to make the answer options that easy that everyone would score 10 out of 10. Likewise, it was possible to make the answer options that difficult that everyone would score 0 out of 10. Yet, the absolute score was nowhere important; it was the relative score that helped to compare employee groups. As said earlier, the polls reflected the strategic priorities for management and were not designed around a previously scientifically validated model.

Secondly, we compared more than 680,000 answers but, at the same time, surveyed only five organizations with five different strategic issues. More comparative studies are needed to see whether the patterns we saw are of a more generic nature across industries and strategic issues. Yet, we also realize that not many organizations want the burden of comparative employee polls, so future researchers need to combine both Likert and Guttman-Poll into the same employee poll.

Thirdly, in terms of investigating the additional insight into employee ambition, there is not enough research yet to determine what combination(s) of ambition Width and –Depth would most benefit an organization. We have used generic assumptions that spending more organizational energy on fewer topics would be beneficial. Hence, more research is needed here.

Further research in larger respondent bodies in various organizations will show whether the patterns and phenomena that emerged in the studied employee polls will materialize elsewhere too.

7. Conclusions

Strategic decision-making is not only about discovering exciting new insights by analyzing large numbers of structured data in corporate data warehouses. Engaging large numbers of employees, if not the input from the entire

organization, results in equally exciting insights. An objective survey scale upgrades the employee poll to a valuable source of data that is a clean 'fuel' to drive algorithms.

Extending this to any organization with strategic issues, this is a fast and affordable way to blend the wisdom of employees in an organization with data already available to management. Moreover, as people have (unlike data in the corporate administration) a rationale and subsequent intentions, a whole new field of analytics lays bare. Similarly, studying more organizations may lead to exploring new patterns that will provide pivotal insights to managers. Not only in strategic decision-making or change management but many other organizational fields such as human resource management, sales and marketing, information technology, and supply chain management as well.

8. References

i. Ahrens, T., & Chapman, C. S. (2006). Doing qualitative field research in management accounting - positioning data to contribute to theory. *Accounting, Organizations and Society, 31*, 819-841. DOI: 10.1016/j.aos.2006.03.007
ii. Conteh, C. (2013). Strategic inter-organizational cooperation in complex environments. *Public Management Review, 15*(4), 501-521. DOI: 10.1080/14719037.2012.674424
iii. De Beuckelaer, A., Weijters, B., &Rutten, A. (2010). Using ad hoc measures for response styles: A cautionary note. Quality & Quantity, 44(4), 761-775. DOI: 10.1007/s11135-009-9225-z
iv. De Jong, M. G., Steenkamp, J. B. E. M., Fox, J. P., & Baumgartner, H. (2008). Using item response theory to measure extreme response style in marketing research: A global investigation. *Journal of Marketing, 45*, 104-115. DOI: 10.1509/jmkr.45.1.104
v. Diamond, I. D., McDonald, J.W., & Shah, I.H. (1986). Proportional hazards models for current status data: application to the study of age at weaning differentials in Pakistan. *Demography 23*(4), 607-620. DOI: 10.2307/2061354
vi. Donaldson, S. I., & Grans-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology 17*(2), 245-260.
vii. Frese, M., & Zapf, D. (1988). Methodological issues in the study of work stress: Objective vs subjective measurement of work stress and the question of longitudinal studies. In: C. L. Cooper, & R. Payne (Eds.), *Causes, Coping, and Consequences of Stress at Work* (pp. 375-411). Wiley & Sons, Chichester.
viii. Giles, J. (2005). Wisdom of the crowd. *Nature 418*(17), 281-281.
ix. Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. 2017. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). 2nd Ed. Thousand Oaks, CA: Sage
x. Hays, R.D., Ellickson, P.L. (1991). Guttman scale analysis of longitudinal data. *The International Journal of the Addictions 25*(11a), 1341-1352.
xi. Koro-Ljungberg, M. (2008). Validity and validation in the making in the context of qualitative research. *Qualitative Health Research, 18*(7), 983-989. DOI: 10.1177/1049732308318039
xii. Kotter, J. (1995). Leading change: why transformation efforts fail. *Harvard Business Review, 73*(2), 59-67. DOI: 10.15358/9783800646159
xiii. Kotter, J. (2012). Accelerate - How the most innovative companies capitalize on today's rapid-fire strategic challenges – and still make their numbers. *Harvard Business Review, 90*(11), 43-58.
xiv. Lake, C. J., Withrow, S., Zickar, M. J., Wood, N. L., Dalal, D. K., &Bochinski, J. (2013). Understanding the relation between attitude involvement and response latitude using item response theory. *Educational and Psychological Measurement*,0013164413482920. DOI: 10.1177/0013164413482920
xv. Likert, R. (1932). A technique for the measurement of attitudes. Archives of psychology.
xvi. Piterenko, K. (2013). Business and impact alignment of questionnaire. *Gjovik University College*.
xvii. Plewis, I., & Mason, P. (2007). What works and why: combining quantitative and qualitative approaches in large-scale evaluations. *International Journal of Social Research Methodology, 8*(3), 185-194. DOI: 10.1080/13645570500154659
xviii. Roulston, K., & Shelton, S. A. (2015). Reconceptualizing Bias in Teaching Qualitative Research Methods. *Qualitative Inquiry, 21*(4), 332-342. DOI: 10.1177/1077800414563803
xix. Guttman, L. (1950). Chs. 2, 3, 6, 8 and 9 in SA Stauffer, L. Guttman, EA Suchman, PF Lazarsfeld, SA Star and JA Clausen. 1950. *Studies in Social Psychology in World War II, 4*, 46-90.
xx. Surowiecki, J. (2005). *The Wisdom of the Crowds.* Anchor, New York. DOI: 10.1119/1.2423042
xxi. Uhlaner, L. M. (2002). The use of the Guttman scale in development of a family business index. (No. H200203). *EIM Business and Policy Research*.

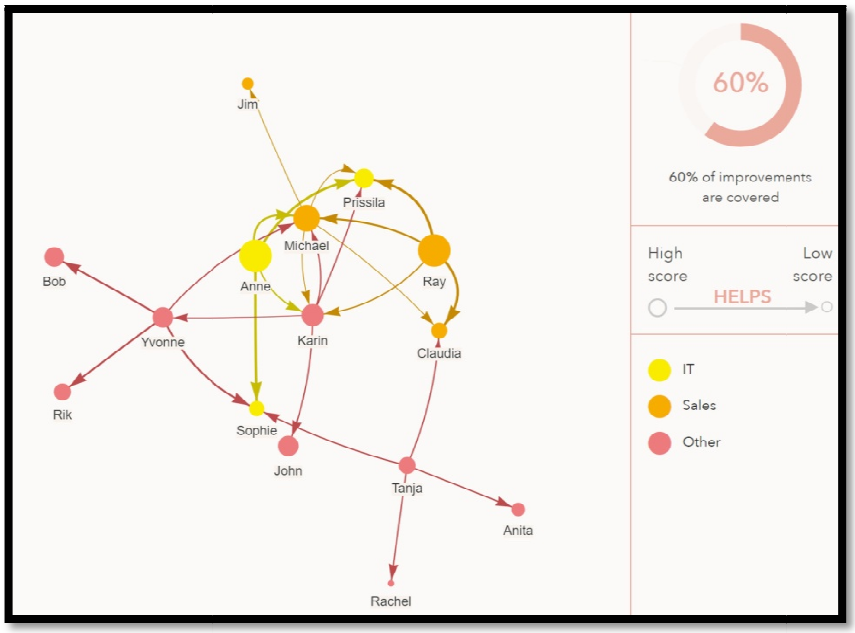

Figure 1: Organizational Network for Knowledge Sharing.