1-1-2008

Method variation in calculating perceived change

Antonis C. Simintiras
Swansea University

Nina Reynolds
University of Wollongong, ninar@uow.edu.au

Method variation in calculating perceived change

Abstract

Motivated by findings in the literature suggesting that error attributed to measures used in generating retrospective reports are excessive, this study explores error attributed to methods that individuals use for calculating change retrospectively. Preliminary findings indicate that method variation is present which, in turn, affects the reported change scores (i.e., the scores varied as a function of the calculation method used). These findings suggest that the accuracy and comparability of retrospective reporting might be improved if one controls for inter-individual calculation method variation. A brief discussion of the implications of the results along with suggestions for future research is provided.

Keywords

change, variation, perceived, method, calculating

Disciplines

Business

Publication Details

Simintiras, A. & Reynolds, N. (2008). Method variation in calculating perceived change. ANZMAC Conference (pp. 1-7). Australia: Australian and New Zealand Marketing Academy.

Method Variation in Calculating Perceived Change

A Simintiras, Swansea University
N Reynolds, University of Bradford

Abstract

Motivated by findings in the literature suggesting that error attributed to measures used in generating retrospective reports are excessive, this study explores error attributed to methods that individuals use for calculating change retrospectively. Preliminary findings indicate that method variation is present which, in turn, affects the reported change scores (i.e., the scores varied as a function of the calculation method used). These findings suggest that the accuracy and comparability of retrospective reporting might be improved if one controls for inter-individual calculation method variation. A brief discussion of the implications of the results along with suggestions for future research is provided.

Introduction

According to Dowling (2001) simple, common-sense ideas of how to measure change in marketing settings can often result in ambiguous and possibly incorrect conclusions being drawn. The overwhelming conclusion is that in contrast to '*before-after*' designs (Byrne and Crombie, 2003), the "use of retrospective accounts in management research needs to be seriously questioned" (Golden, 1992:857). However, given a) the limited research into the intricacies of measuring change retrospectively, b) the popularity of retrospective reports, and c) findings suggesting that error attributed to informant fallibility is not excessive but error attributed to measures used in generating the retrospective reports is excessive (Miller *et al.,* 1997), this study examines whether or not there is calculation method variation that could account for error in retrospective reports of perceived change.

The purpose of this study is to examine the method(s) that individuals use to calculate change retrospectively for successive past time periods. The paper provides a brief review of the concept of change and its measurement, and presents an approach for identifying and describing different calculation methods. The paper concludes with a discussion of the benefits accruing from the findings in the light of the study's limitations.

The Measurement of Change

The basic approach to measuring change is to calculate the simple difference between multiple measures of the same variable over time. For example change of variable X of a person *j* (*j = 1, 2, 3..., n*) over two time periods can be expressed as: $C_x = X_{1j} - X_{2j}$, where C_x is the change score, X_{1j} is the variable X measured in time period 1, and X_{2j} is the variable X measured at time period 2.

The main research designs for gathering change data are longitudinal and cross-sectional. In longitudinal - as in '*pretest-posttest'* - designs the pretest is administered before the intervention (treatment) and the change is measured as the difference between pretest and posttest scores (Goldstein, 1968). In cross-sectional designs, '*retrospective pretests*' are administered at the end of the intervention (treatment), at the same time and often on the same form as posttest ratings (Hill and Betz, 2005). Usually, respondents are instructed to think about and complete the retrospective ratings first and then to complete their ratings as

they apply at present. The change is assessed as the difference between '*thentest*' and '*posttest*' scores. Another design for measuring change retrospectively is the *'perceived change'* where respondents are instructed to report perceived change, which may (or may not) be the result of an intervention (Lam and Bengo, 2003).

Validity Concerns

Data from longitudinal and cross-sectional designs have inherent weaknesses as far as their validity is concerned (Hilton and Patrick, 1970). In longitudinal studies, (e.g., *pretest-posttest*) when a sample of individuals is tested for the second time and a change is observed, it is difficult to establish whether the observed change is *alpha*, *beta* or *gamma* type change (Golembiewski *et al.,* 1976). "*Alpha* change – occurs when the meaning of the construct to the respondent and the psychological interpretation of the units of measurement on the 'measurement rule' stays the same, but the level of the measurement on this rule changes; *beta* change – arises where the meaning of the construct to the respondent stays the same, but the respondent subjectively recalibrates the 'measurement rule'; and g*amma* change – occurs where the conceptual domain of the construct … is radically altered in such a way as to make the previous meaning of the 'measurement rule' irrelevant." (Dowling, 2001:56). Consequently, the observed changes may be solely the result of the intervention (e.g., *alpha change*), or they may be due to error (e.g., *beta change* and/or *gamma change*).

In cross sectional designs for measuring change retrospectively (e.g., *'thentest-posttest'* and *'perceived change'*) the overwhelming evidence is that these measurements suffer from multiple forms of bias. According to Smith (1984), there are notable problems relating to forgetting and memory distortion and the three most common errors are: (1) *Forgetting* (see Beckett *et al.,* 2001); (2) *Time displacement* (see Huttenlocher *et al.,* 1990); and (3) *Distortion* (see Beckett *et al.,* 2001). Errors are also caused by imperfect recall capabilities (see Erickson and Simon, 1980), the influence of the Law of Small Numbers and the accessibility principle (see Brown *et al.,* 1986), the over-response to vivid information (see Huber and Power, 1985) and the hindsight effect (see Fischhoff, 1982). The hindsight bias leads individuals to "consistently exaggerate what could have been anticipated in foresight." (Fischhoff, 1982:341). Another source of error, is the possibility of inaccurate reporting due to a conscious misrepresentation (Sudman and Bradburn, 1974) in order to be more congruent with self perceptions and an attempt to present an image more in line with what they feel to have occurred or to give a more 'socially desirable' response (Powers *et al.,* 1978).

Proponents of *'pretest-posttest'* argue that the validity concerns inherent in retrospective data result in unacceptable levels of measurement bias; proponents of *'thentest-posttest'* argue that response shift bias (i.e., gamma change) poses a greater problem than self report bias. Both sets of validity concerns result in the same prediction; namely, the change scores calculated from true pretest scores will be smaller than those calculated from retrospective pretest scores (Norman, 2003). Those who favor the '*pretest-posttest*' designs and those who favor the '*thentest-posttest*' designs, however, interpret the discrepancy between the change scores, differently. Hill and Betz (2005) concluded that replacing traditional (*pretest-posttest*) with retrospective pretests (*thentest)* does not eliminate bias, and recommended the use of *pretest-posttest* for examining the intervention effects and '*thentest-posttest*' for investigating subjective experiences of intervention-related change.

Tompkins and Cheney, (1983), stated that it is doubtful that an accurate representation of most types of actions or behaviors can be produced, and Singer (1977) argued that there is no

way to tell how accurately subjective appraisals of change reflect objective conditions; often, these appraisals appear to reflect a comparison between what is and what was expected, rather than a comparison between what is and what was. Therefore, subjectively reported change, is not a good substitute for measures of change derived from '*before-after*' designs though, at times, it may be the only alternative (Singer, 1977).

Measuring change retrospectively calls for an understanding between 'perceived change' and 'actual change'. Specifically, *'perceived change'* ought to be concerned with the experienced change (i.e., perceived magnitude) rather than the actual change (i.e., absolute magnitude). On the usefulness of measuring *'perceived change'*, Blane (1996) has argued that careful examination of issues pertaining to 'which items of information can be recalled with what degree of accuracy' and 'how accuracy can be improved' constitutes a more constructive approach (Berney and Blane, 1997) rather than the ad-hoc rejection of this type of data as invalid. Consequently, the method individuals use to calculate change in retrospective studies becomes an important issue as it impacts on the accuracy of reporting; this study, investigates only the method variation in calculating change without taking into consideration the accuracy of the recall of past events.

Methodology

In order to investigate whether individuals use different methods for calculating perceived change retrospectively, a sample of MBA students (n=30) was asked to complete a simple exercise. Initially, students were asked to choose an activity that they were engaged in the past. The range of activities provided was: sports, entertainment and shopping. Next, respondents were asked to recall and report the number of hours spent on average per week on the chosen activity in year 2002 and then recall and report activity change for each of the following three consecutive years (i.e., 2003, 2004 and 2005). In addition, subjects were asked to a) calculate the change in the number of hours (in percentages) for each time period, b) estimate the overall change (increase, decrease or no change) in the number of hours for the entire period (2002 to 2005), and c) calculate the percentage change for the entire period (2000-2005). Specific instructions were given to subjects prior to the exercise and they were allowed the time needed to recall, calculate, and report activity change (Threlfall, 2002). Furthermore, the subjects were asked to provide written information regarding the method they used to calculate percentage change for each year and for the entire period. This approach is based on Barnett and Carroll's (1995) suggestion that change to be reported should be repeatable and hence comparable across repetitions. Three waves of pre-tests carried out before the final version of the exercise (questionnaire) was finalized. Subjects used for the pretests were postgraduate students, other than MBAs, and academic staff.

Preliminary Findings

The breakdown of the chosen activities was as follows: sports (n=14), shopping (n=10), and entertainment (n=6). The sample comprised 19 males and 11 females from 13 different countries. A large number of respondents (n=13) returned non-usable answers - five subjects failed to report change for more than one period, whereas the calculations of the other 8 respondents were numerically inconsistent. The remaining 17 subjects provided logical consistent calculations and their answers were examined further to identify the calculation methods that they used. Four clearly identifiable methods were found. These are: Initial-base, Re-base, Cumulative, and Adjustment. The usage frequency of each method is shown in Table 1:

Table 1: Frequency of calculation methods used per chosen activity

	Initial-base	Re-base	Cumulative	Adjustment	Total
Sports	3	3	0	2	**8**
Entertainment	1	2	0	1	**4**
Shopping	2	2	1	0	**5**
Total	**6**	**7**	**1**	**3**	**17**

The most frequently used methods were the 'Re-base' and the 'Initial-base' (7 and 6 respondents respectively). Two of the respondents used more than one method. One respondent used the re-base method for calculating change for the first two years and the adjustment method for the third year, whereas the other respondent used the re-base method for the first two years and the initial-base method for the third year.

Description of Methods for Calculating *'Perceived Change'*

For illustrative purposes and consistency in presenting the methods, it is assumed that the initial base at t_{2002} was 10 hours on average per week and that a respondent recalled the following increases in number of hours (average per week) for the chosen activity: For $t_{2003,}$ = 4 hours, for $t_{2004,}$ = 2 hours, and for t_{2005} = 3 hours. The percentage change reported for each method is shown in Table 2; how these percentages are reached is explained below.

Table 2: Percentage change reported in hours of activity

	Hours of Activity	Percentage change reported			
		Initial-base	Re-base	Cumulative	Adjustment
t_{2002}	10	---	---	---	---
t_{2003}	14	40%	40%	40%	40%
t_{2004}	16	20%	14%	60%	50%
t_{2005}	19	30%	19%	90%	150%
"Total" change		**90%**	**73%**	**190%**	**240%**

Initial-base method: Respondents using the initial-base method treated the first time period (e.g., year) as a base for subsequent calculations. The reported percentage change for each time period was calculated using the initial base (number of hours spent on activity at t_{2002} for each time period (e.g., $t_{2003,}$ t_{2004}, and t_{2005}). For example:

t_{2003} = (4/10) x 100 = 40%; t_{2004} = (2/10) x 100 = 20%; t_{2005} = (3/10) x100 = 30%.

Overall perceived change = 40% + 20% + 30% = *90%* (or 9/10 = 90%).

Re-base method: Respondents using the re-base method treated each time period (e.g., year) as a new base for calculating change for the next period. The percentage change for a period was based on the total number of hours calculated for the previous period. For example:

t_{2003} = (4/10) x100 = 40%; t_{2004} = (2/14) x100 = 14%; t_{2005} = (3/16) x100 = 19%.

Overall perceived change = 40% + 14% + 19% = *73%* (or 4/10+2/14+3/16 = 409/560 = 73%)

Cumulative method: The cumulative method was used by one respondent, who calculated change cumulatively for each successive time period. Consequently, the change for one period depended on the change of the previous period (change for a preceding period was combined with the change of the new period under consideration cumulatively). For example:

t_{2003} = (4/10) x100 = 40%; t_{2004} = (6/10) x100 = 60% (e.g., 4 hours for the first period + 2 hours for the second period); t_{2005} = (9/10) x100 = 90% (e.g., 4 + 3+ 2 = 9).
Overall perceived change = 40% + 60% + 90% = *190%* (or 19/10 = 190%).

Adjustment method: Three respondents used the adjustment method. They calculated activity change relative to the change of the preceding period. More specifically, respondents started from the initial base and proceeded to estimate percentage change for each successive period using as a base the change in the previous time period. For example:

t_{2003} = (4/10) x100 = 40%; t_{2004} = (2/4) x100 = 50% (e.g., 2 hour change in the second time period relative to a 4 hour change for the first time period); t_{2005} = (3/2) x100 = 150% (e.g., 3 hour change in the third time period relative to a 2 hour change for the second time period).
Overall perceived change = 40% + 50% + 150% = *240%* (or 4/10+2/4+3/2 =24/10 = 240%).

The above findings indicate that inter and intra individual calculation method variation exist and this variation results in different scores in perceived change.

Discussion

This study informs researchers of the impact of method variation on the comparability of retrospective measures of change, though the small sample size and the large number of non-usable responses is a cause of concern. Establishing whether or not the differences in retrospective measurements are due to inaccurate and biased recall of past events (e.g., memory failure, telescoping effects) or method variation is important. The focus of the existing literature is on recall bias as opposed to error due to method variation in calculating change. From a practical perspective, scores resulting from inter-individual method variation can be recalibrated and compared, but only when the researcher has *a-priori* information of the method used by each respondent. In that case, the problem of method variation might be contained to the level of respondents' ability to use the chosen method correctly and consistently.

Intra-individual method variation was also found to affect the reported change scores; this poses an additional threat to the interpretation of change scores. The critical issue, however, is whether or not intra-individual method variation influence goes beyond the reported scores to cause distortions in memory as, perhaps, reflected by the disaggregating mechanism in the process of estimating perceived change. Inter-individual calculation method variation influences the comparability of scores. Intra-individual method variation influences go beyond memory and telescoping effects to the structuring and manifestation of the process of recall itself (i.e., accuracy). Both are areas in need of further research.

The range of methods identified in this study is not exhaustive. The findings suggest that there is an influence of calculation method used on the reported scores of perceived percentage change. Also, from the numerically inconsistent results of eight of the excluded respondents, there is tentative evidence suggesting that individuals alternate between methods causing intra-individual method variation. Although the findings can be considered by researchers measuring change retrospectively, the entire area requires a more comprehensive analysis and a more holistic approach.

References

Barnett W.P., Carroll, G.R., 1995. Modeling internal organizational change. Annual Review of Sociology 21, 217-236.

Beckett, M., DaVanzo, J., Sastry, N., Panis, C., Peterson, C., 2001. The quality of retrospective data: An examination of long-term recall in a developing country. The Journal of Human Resources 36(3), 593-625.

Berney, L.R., Blane, B.D., 1997. Collecting Retrospective data: Accuracy of Recall after 50 years judged against historical records. Social Sciences and Medicine 45(10), 1519-1525.

Blane, B.D., 1996. Collecting retrospective data: Development of a reliable method and a pilot study of its use. Social Science and Medicine 42, 751-757.

Brown, N.R., Rips, L.J., Shevell, S.K., 1986. The subjective dates of natural events in very long term memory. Cognitive Psychology 17, 139-177.

Byrne, M.B., Crombie, G., 2003. Modeling and testing change: An introduction to the latent growth curve model. Understanding Statistics 2(3), 177-203.

Dowling, G.R., 2001. The alpha, beta, gamma approaches to measuring change and its use for interpreting the effectiveness of service quality programs. Australian Journal of Management 26(1), 55-68.

Ericsson, K.A., Simon, H.A., 1980. Verbal reports as data. Psychological Review 87, 215-251.

Fischhoff, B., 1982. For those condemned to study the past: Heuristics and biases in hindsight. In: Kahneman, D., Slovic P., Tversky A., (Eds.), Judgment under uncertainty: Heuristics and biases. Cambridge University Press, Cambridge UK

Golden, B.R., 1992. The past is the past – or is it? The use of retrospective accounts as indicators of past strategy. Academy of Management Journal 35, 848-860.

Goldstein, H., 1968. Longitudinal studies and the measurement of change. The Statistician 18(2), 93-117

Golembiewski, R., Billingsley, K., Yeager, S.M., 1976. Measuring change and persistence in human affairs: Types of change generated by OD designs. Journal of Applied Behavioral Science 12, 133-157.

Hill G.L., Betz, L.D., 2005. Revisiting the retrospective pretest. American Journal of Evaluation 26(4), 501-517.

Hilton, T.L., Patrick, C., 1970. Cross-sectional versus longitudinal data: An empirical comparison of mean differences in academic growth. Journal of Educational Measurement 7(1), 15-24

Huber, G.P., Power, D.J., 1985. Retrospective reports of strategic-level managers: Guidelines for increasing their accuracy. Strategic Management Journal 6(2), 171-180.

Huttenlocher, J., Hedges, L.V., Bradburn, N.M., 1990. Reports of elapsed time: Bounding and rounding processes in estimation. Journal of Experimental Psychology, Learning, Memory, and Cognition 16, 196-213.

Lam, T.C.M., Bengo, P., 2003. A comparison of three retrospective self-reporting methods for measuring change in instructional practice. American Journal of Evaluation 24, 65-80.

Miller, C.C., Cardinal, L.B., Glick, W.H., 1997. Retrospective reports in organizational research: A re-examination of recent evidence. The Academy of Management Journal 40(1), 189-204.

Norman, G., 2003. Hi! How are you? Response shift, implicit theories and differing epistemologies. Quality of Life Research, 12(3), 239-249.

Powers, A.E., Goudy, J.W., Keith, M.P., 1978. Conguence between panel and recall data in longitudinal research. Public Opinion Quarterly 42, 380-389.

Singer, E., 1977. Subjective evaluations as indicators of change. Journal of Health and Social Behavior 18, 84-90

Smith, W.T., 1984. Recalling attitudes: An analysis of retrospective questions on the 1982 GSS. Public Opinion Quarterly 48, 639-649.

Sudman, S., Bradburn N.M., 1974. Response effect in surveys: A review and synthesis. Aldine, Chicago.

Threlfall, J., 2002. Flexible mental calculation. Educational Studies in Mathematics 50(1), 29-47.

Tompkins, P.K., Cheney, G., 1983. Account analysis of organizations: decision making and identification. In: Putnam, L., Pacanowsky, M.E., (Eds.), Communication and Organizations: An Interpretive Approach. Sage, Beverly Hills, CA.