Optimizing Survey Measurement Accuracy by Matching Question Design to Respondent Memory Organization

Michael D. Silver and Jon A. Krosnick
Ohio State University

Abstract

Surveys often ask people to report the frequency with which they have performed various types of behaviors or witnessed various types of events. Although past research has shown many ways in which the ordering of such questions can bias frequency reports, we know little about how to optimize the ordering of such questions to maximize the accuracy of people's reports. This research took as its starting point the notion prevalent in the literatures of social and cognitive psychology that pieces of knowledge are organized in associative networks that cluster related information together. Therefore, optimal measurement may occur when questions addressing knowledge that is stored in a memory cluster are asked as a group before moving on to asking questions addressing knowledge stored in a different cluster. We report the results of four studies, using a variety of techniques to diagnose the memory organizations of commercial airline pilots, lending support to this general proposition.

Authors' Note: The authors thank Robert Dodd, Loren Rosenthal, and Linda Connell for making this research possible, Mike Jobanek, Augustina Jay, Alodia Velasco, Michael Walker, and Reid Hastie, who provided help and advice, and the many pilots who participated in these studies.

Keywords: Questionnaire design, validity, memory organization, factual recall

Introduction

Surveys routinely ask people to report the frequency with which they have performed various behaviors or witnessed various sorts of events in the past. In order to design a questionnaire for such a survey, items must be written asking about each type of behavior or event, and then those items must be placed in a particular order for administration. In the research described in this paper, we explored whether recall accuracy can be optimized by ordering such questions according to the organization of relevant information in respondents' memories.

The inspiration for this work came from the literature on memory organization in psychology. According to this literature, it is useful to think of memories of individual events as organized in associative networks (e.g., Collins & Loftus, 1975), which link related memories to one another in a category. Memories of events that are dissimilar and unrelated are unlikely to be interlinked or to form a category.

Recalling a memory from one category generally makes it easier to recall other memories from that same category subsequently. For example, if a person were asked to recall one event in the category "dining out," doing so would make it easier for him or her to recall other dining out events if they are clustered together in his or her memory. By the same token, recalling an experience of dining out would probably not make it easier to recall instances of oil changes and might even interfere with such recall, because irrelevant memories are activated and may be distracting. This logic implies that factual recall in surveys might be improved by asking questions about events that are associated within one mental category before moving on to ask questions in another category.

In the studies we describe here, we tested several possible organization schemes that we thought might represent memory organizations for specific events. Then, once we identified a plausible memory organization, we gauged the effect of using it in survey question design on the accuracy of recollections of events.

We conducted this research as part of a project sponsored by the National Aviation and Space Administration (NASA) to build a survey questionnaire that would measure the frequency with which commercial airline pilots experienced aviation safety-related events (e.g., blocked radio transmissions, severe turbulence, near mid-air collisions, etc.). A priori, we identified three possible schemes by which pilots might organize such events in their memories: in terms of their causes (e.g., equipment failures, flight crew mistakes, weather, air traffic controllers, passengers), the phase of flight when they occur (e.g., take-off, cruise, landing), and their seriousness (minimal, moderate, or severe). Rather than choose one of these potential organization schemes for our survey questionnaire, we tested the effectiveness of the schemes at describing pilots' memory organizations and for improving recall accuracy. The four studies reported here describe that process and the results.

A diverse set of 68 commercial pilots with a wide range of experience participated in the four studies we describe here. In Study 1, pilots were asked to remember all the aviation safety related events they had experienced during their careers. We examined the order in which pilots remembered events and how they labeled groups of related events. In Study 2, pilots sorted 96 fictitious safety-related events into groups, indicating which events seemed most similar or related to one another. Analyses of how the pilots labeled the stacks and of how the events were clustered within stacks indicated how the pilots organized the events. In Study 3, pilots read descriptions of the 96 events and later tried to remember them. We again examined the order in which pilots remembered events and how they labeled groups of related events. Studies 1, 2, and 3 identified an unexpected organizational scheme that appeared to be most commonly used by pilots, which we called the "hybrid" scheme. Study 4 experimentally gauged whether a questionnaire utilizing the hybrid scheme in fact yielded improved recall as compared to questionnaires using other schemes.

Exploratory Analyses of the Autobiography, Sorting, and Recall Data

In this section, we describe the data collection methods and exploratory analysis results for Studies 1-3.

Study 1: Autobiographies. Nine pilots were asked to describe all of the air safety problems they had personally witnessed during their flying careers. Responses were tape-recorded for later transcription and analysis. Pilots mentioned both specific events (e.g., "One time, I nearly went onto the wrong runway") and more general categories of events (e.g., "altitude deviations"). The first column of Table 1 displays most of the verbatim general category statements pilots made.[1] The general categories mentioned were combined into what we call event clusters, based upon the NASA ASRS (Aviation Safety Reporting System) organizing scheme and our a priori expectations. The second column of Table 1 lists the clusters.

[1] If a pilot mentioned two or more general categories that were ultimately coded in the same event cluster in column 2 of Table 1, only one of these general categories appears in column 1.

Our interest was in identifying clusters that were used by most pilots in organizing safety-related events in memory. Therefore, the third column of Table 1 shows the percent of pilots who mentioned events in each cluster at least once. Over 50% of the pilots mentioned equipment-, cockpit crew-, and ATC-related events. These most commonly-mentioned clusters all focus on causes, which suggests that causes may be a primary organizational principle used by pilots. Reinforcing this notion is the prevalence of causes among the clusters mentioned by 22% of pilots (passengers, ground personnel, weather, and turbulence) and among those mentioned by 11% of pilots (flight attendants). Among the less-often mentioned clusters, some are phases of flight (e.g., preflight, taxiing, approach), and others refer to severity (minor problems, moderate problems). Severity categories were mentioned by only single pilots, whereas some phases of flight categories were mentioned by two pilots, which suggests that the latter may be more commonly used than the former. But it is clear that neither were used very widely.

Study 2: Event Sorting. Fourteen pilots were each given a stack of 96 index cards, with a one- or two-sentence description of a fictitious safety-compromising event printed on each one. The cards were presented to each pilot in a unique random order. Each pilot sorted the cards into as many stacks as he or she felt was appropriate to indicate which events had something in common and went together as a group. When a pilot completed sorting the cards, he or she was asked to label each stack "to indicate what the cards in the stack have in common."

Each event fell in one cell in an 8 (flight phase: preflight, taxiing in and out, take-off, climb, cruise, descent, approach, and landing) x 4 (cause: equipment, weather, people, and other vehicles) x 3 (severity: no adverse results, moderate adverse results, and severe adverse results) matrix; each event represented a particular combination of the three attributes (see Table 2). An example is: "While starting engines for a flight from Boston/Logan (BOS) to Baltimore (BWI), a crew experienced a passenger-initiated aircraft evacuation. A passenger yelled "fire" during engine start. Frightened passengers opened over-wing hatches and aft exits and began exiting the aircraft. Several passengers were injured" (Cause: People, Phase: Pre-flight, Severity: Severe).

The labels that pilots wrote to describe the stacks were again grouped into clusters, which are listed in the second column of Table 2. For example, 79% of pilots used at least one of the three labels in the "equipment" cluster in labeling one of their stacks (see the first row of the table).

A comparison of Studies 1 and 2 (see columns 1 and 2 of Table 2) reveals some correspondence in the frequency of the clusters. In both studies, the three most frequently mentioned clusters were causes (e.g., equipment problems, cockpit crew problems), and equipment problems were mentioned most often in both studies. However, some clusters related to the location of the events (i.e., mid-air conflicts and events on the ground) were mentioned by over 50% of the pilots in the sorting study, yet they were mentioned much less frequently in the autobiography study. Additionally, some severity-related clusters (e.g., minor problems) were among those mentioned by 50% of the pilots in the sorting study, although no such clusters were mentioned frequently in the autobiography study. As in Study 1, causes appeared to be widely-used organizing clusters. However, other sorts of clusters appeared to be widely used as well. But we saw no indication from these labeling data that either causes or locations or phases or severity completely dominate all others as organizing principles.

Study 3: Event Recall. In Study 3, nine pilots read the 96 event descriptions (in a unique random order for each pilot), performed a ten-minute “distraction task,” and then tried to recall the events as accurately as possible. The pilots’ statements were tape-recorded and later transcribed.

Table 1: General Categories of Events from Pilots’ Autobiographies in Study 1

General Categories	Clusters Derived from Pilots’ Statements	Percentage of Pilots Using Cluster at Least Once
Engine shutdown. Air conditioning packs…out of service. Erroneous stall warning indications. Maintenance things. Maintenance. Air crew…sign off maintenance problems that are not life threatening.	Equipment	67%
Pilot error. I've … spun the wrong knob. Conversations in the cockpit that happen during the critical phases of flight. Cockpit talk…they violate our sterile cockpit rule. Discipline in the cockpits. It's clear as a bell to us what we thought we heard.	Cockpit Crew	67%
Vectored to final very high, very fast. Getting clearance readbacks. Common terminology. Misunderstandings (between ATC and crew). Interaction between controllers and pilots.	Air Traffic Controllers	56%
Mid-air collision potential. We have mid-airs or near mid-airs. Near miss situations.	Mid-Air Conflicts	33%
Passengers in the vicinity of the propellers. Medical problems.	Passengers	22%
Ramp safety, people moving around the aircraft. Ramp people.	Ground Personnel	22%
Varying weather conditions. Icy conditions .	Weather	22%
Wake turbulence. Caught in the other plane's wake.	Turbulence	22%
When we're taxiing. Rolling on the taxiway.	Taxiing	22%
Approach phase. Go-arounds.	Approach	22%
A flight attendant…stays a little longer (in the cockpit) than he or she should.	Flight Attendants	11%
When we're loading bags	Preflight	11%
Just after rotation.	Takeoff	11%
An ultra minor thing	Minor Problems	11%
They don't …cause a big threat to safety, but probably some threat	Moderate Problems	11%
A big safety area.	Severe Problems	11%

Again, pilots sometimes named general categories of events (e.g. "Well, there were some aborts"). The third column of Table 2 lists the clusters we derived from those statements. Some aspects of these results correspond to those in columns 1 and 2 of Table 2. In all three studies, a cause cluster was the most frequently mentioned. Two other cause clusters, passenger problems and weather problems, were mentioned by 67% or more pilots in the sorting and recall studies. As in the sorting study, more than 50% of the pilots in the recall study mentioned some location clusters (i.e., mid-air conflicts and events on the ground). And as in the sorting study, severity clusters were rarely mentioned in the recall study. In general, causes were widely mentioned by pilots in the recall study, and locations were mentioned somewhat as well.

Creating a Hybrid Organizational Scheme

To best ascertain the prevalence of various clusters across pilots, we collapsed the data obtained from Studies 1-3, as shown in Table 2. Column 1 displays all the clusters that emerged in those studies. Columns 2, 3, and 4 display the proportions of pilots who mentioned each cluster in each study. Column 5 displays the percent of all pilots who mentioned that cluster at least once. Table 2 is divided vertically into three sections, differentiating clusters mentioned very frequently (by more than one-third of respondents), moderately frequently (by more than one pilot but fewer than one-third of them), and rarely (by only one single pilot). Most of the clusters in the top section of the table were mentioned by at least one pilot in each one of the three studies, reinforcing the notion that they are widely used.

Causes clearly dominate this set, including equipment, cockpit crew, passengers, weather, turbulence, and ATC. In fact, the three most frequently mentioned clusters were all causes. However, other sorts of clusters were also present among this most-prevalent set. Specifically, two location clusters (mid-air conflicts and events on the ground), one severity cluster (injuries), and one flight phase cluster (taxiing) were mentioned by more than one-third of respondents. In addition, altitude deviations were mentioned frequently, this cluster label referring to a type of event, rather than a specific cause, flight phase, or severity. Consequently, the most common clusters are not all of one type, which suggests that no single organization scheme among those we identified a priori (cause, flight phase, or severity) would be maximally efficient in facilitating pilots' recall of safety-related events.

The results of Study 2 correlate fairly strongly with those of Study 1, $r(36)=.63$, $p<.001$, and with those of Study 3, $r(36)=.67$, $p<.001$. The results of Studies 1 and 3 correlate less strongly, $r(36)=.32$, $p=.06$.

More importantly, the results of each study correlated strongly with the results obtained when all three studies' findings were combined (shown in the last column of Table 2). The correlations between the combined frequencies and the frequencies from the three individual studies were: .94 for the sorting study, .84 for the recall study, and .70 for the autobiography study (all p's$<.001$). This extremely high level of correspondence justifies confidence in the validity and usefulness of the high-frequency clusters identified in the last column of Table 2.

Based upon these results, we developed a hybrid scheme capturing the clusters used most often, such that each item in a proposed questionnaire would fall unambiguously in one cluster. In resolving overlap among clusters, we placed priority on clusters mentioned by more pilots. The resulting scheme is: <u>Equipment</u>: Any equipment problem; <u>Cockpit Crew</u>: Any cockpit crew-

Table 2: Clusters Derived From Pilots' Statements in Studies 1, 2, and 3

	Study			
Cluster from Pilots' Statements	Study 1: Autobiography	Study 2: Sorting	Study 3: Recall	Studies 1-3 Combined
Frequently Mentioned				
Equipment	67%	79%	78%	75%
Cockpit Crew	67%	71%	33%	59%
Passengers	22%	79%	67%	59%
Mid-Air Conflicts	33%	57%	89%	59%
Weather	22%	71%	67%	56%
Turbulence	22%	36%	100%	50%
Events on the Ground	0%	57%	89%	50%
Air Traffic Control	56%	50%	0%	38%
Injuries	0%	50%	56%	38%
Altitude Deviations	0%	29%	78%	34%
Taxiing	22%	43%	33%	34%
Moderately Frequently Mentioned				
No Injuries	0%	43%	44%	31%
Takeoff	11%	29%	44%	28%
Landing	0%	36%	33%	25%
Minor Problems	11%	50%	0%	25%
Approach	22%	21%	22%	22%
Severe Problems	11%	21%	22%	19%
Flight Attendants	11%	0%	44%	16%
Preflight	22%	14%	22%	16%
Climb	0%	14%	22%	13%
Aircraft Damage	0%	29%	0%	13%
Descent	0%	14%	11%	9%
Cruise	0%	7%	11%	6%
Enroute	0%	7%	11%	6%
Moderate Problems	11%	7%	0%	6%
Ground Personnel	22%	0%	0%	6%
Crashes	0%	14%	0%	6%
Rarely Mentioned				
Departure	0%	0%	11%	3%
Fire/Smoke	0%	7%	0%	3%
Controlled Flight Into Terrain	0%	7%	0%	3%
Planning.	0%	7%	0%	3%
Rapid depressurization.	0%	7%	0%	3%
Military.	0%	7%	0%	3%
Construction.	0%	7%	0%	3%
Could fix or hurt you.	0%	7%	0%	3%
Close calls.	0%	7%	0%	3%

Note. Cell entries are the percentages of pilots who mentioned a general category in each cluster at least once.

caused problems/errors not resulting in an air conflict or ground problem; Passengers: Any passenger-caused problems; Mid-Air Conflicts: Any conflicts with other aircraft in the air; Weather: Weather problems other than turbulence; Turbulence: Any turbulence, including wake turbulence; Events on the Ground: Any runway or taxiway transgressions and any ground conflicts with other vehicles; Altitude Deviations: Any altitude deviations not resulting in a conflict. This scheme includes all of the clusters mentioned by more than one-third of respondents across the first three studies, except for three: ATC (air traffic controllers), injuries, and taxiing.[2]

To assure that the hybrid scheme's clusters were defined in a way that minimized ambiguity, we imposed a hierarchical organization on the clusters based upon the frequency with which each cluster was mentioned across our three studies. An event that could legitimately be placed into more than one cluster was placed in the one earlier in the listing above.

Confirmatory Testing of Pilots' Memory Structure

The data collected in Studies 2 and 3 afforded opportunities to test the viability of the hybrid scheme in comparison with the three other schemes that we thought a priori might be used by pilots (cause, flight phase, and severity).

Cluster Analysis of Stacks in Study 2. The manner by which pilots organized the safety-related events into stacks in Study 2 can provide an indication of how safety-related events are organized in their memories. One method for identifying the organization of memory using sorting data is cluster analysis (Brewer & Lui, 1996). Using a confirmatory method of cluster analysis (Arabie & Carroll, 1980), we compared the fit of the observed clustering to the clustering predicted by each of the four proposed organizational schemes.

To assess the co-occurrence of events in stacks in the sorting study, we created a 96x96 co-occurrence matrix for each pilot for the 96 statements he or she sorted. The aggregation of these matrices was submitted to the computer program STANDAL (Walker, 1998) along with codes indicating the organizational clusters into which each of the 96 events fell. The program converted the matrix to a format that could be analyzed using standard multiple regression in a procedure known as stand-alone regression (Arabie & Carroll, 1980).

This was done four separate times, once for each of the four contending organizational schemes (hybrid, cause, flight phase, severity). For all four analyses, we entered into the program the same observed co-occurrence matrix. But for each of the four organizational schemes, we gave the program a unique set of codes that matched each of the 96 events to each cluster (e.g., take-off, etc.) within that potential organization (i.e., flight phase).

---

[2]The "injuries" category was eliminated because events falling in that category could fall into all the others as well. Taxiing was eliminated because events falling into that category would presumably fall into the category of events on the ground as well, the latter having been mentioned more frequently than the former. And air traffic controllers was eliminated because the problems they create probably most often yield mid-air conflicts or events on the ground.

The STANDAL program produced a new data file, each row of which corresponded to one of the pairs of events in the 96x96 matrix. The first variable represented the observed co-occurrence of the two events in the same stack (standardized from the original co-occurrence matrix to range from zero, meaning no co-occurrence, to one, meaning co-occurrence for all 14 pilots). Also in each row were a series of dummy variables created by STANDAL to represent how the items were hypothesized to cluster. Each of the other variables represented a different cluster (e.g., take-off, cruise), and the value for that variable indicated whether the events were hypothesized to fall together in that cluster (coded 1) or not (coded 0).

Once a data file for an organizational scheme was constructed in this way, an ordinary least squares multiple regression was conducted, predicting the observed co-occurrences from the hypothesized co-occurrences (represented in the dummy variables) for that organizational scheme. The regression results provided estimates of how well each scheme accounted for how the pilots sorted the events. In essence, we tested whether events thought to fall into the same cluster were in fact sorted together more often than events thought to fall into different clusters.

One regression was performed for each of the four hypothesized organizational to calculate the proportion of variance in the observed clustering accounted for by each hypothesized structure. The flight phase and severity schemes accounted for very little variance: 5% and 1%, respectively. The cause scheme accounted for much more variance: 26%. But the best fit came from the hybrid structure, which explained a 34%. The hybrid scheme explained significantly more variance in the sorting data than did the flight phase scheme, $F(8,4543)=248.67$, $p<.001$, the cause scheme, $F(8,4547)=73.20$, $p<.001$, and the severity scheme, $F(8,4548)=280.82$, $p<.001$.

Cluster Analysis of Event Recall Order in Study 3. Cognitive organization of events in memory can also be seen in the recall study by analyzing the order in which specific events were recalled. When a pilot recalls one event, he or she is then more likely to recall another event from that same cluster than an event from a different cluster. Using this logic, we compared the order in which a pilot recalled events with the order in which we would expect the events to be recalled according to each of the four schemes.

For each pilot, we computed an Adjusted Ratio of Clustering or ARC score (Roenker, Thompson, & Brown, 1971) for each of the four hypothesized organizational schemes. ARC scores of 1 mean that all events recalled by a pilot thought to fall in the same cluster were in fact recalled sequentially. An ARC score of zero means that knowing which hypothesized clusters events fall in provides no basis for predicting the order in which a pilot recalled events. (A negative ARC score would mean that a pilot tended to alternate between hypothesized clusters in recalling events, rather than recalling more than one event from a hypothesized cluster sequentially.)

We computed an ARC score for each organizational scheme for each pilot. We then conducted a repeated-measures analysis of variance on the ARC scores of all nine pilots to test for differences in the amounts of clustering in the recall data (see Ostrom, Carpenter, Sedikides, & Li, 1993). None of the mean ARC scores for the three a priori schemes differed significantly from zero, meaning that these schemes did no better than chance at predicting recall order (Flight phases: mean=.11, p=.47; Causes: mean=-.01, p=.94; Severity: mean=.02, p=.81). As expected, however, the hybrid scheme's mean ARC score was significantly positive (mean=.30, p=.04), meaning that it did significantly better than chance at accounting for the observed recall order.

Study 4: The Experiment. If the hybrid organizational scheme fits most pilots' memories best, then employing it in a questionnaire design should facilitate pilots' recollection of events they have witnessed. To test this claim, we conducted an experiment in which 36 pilots read the 96 fictitious events and later attempted to recall them. Some did so with no help, just as participants in Study 3 had earlier. Other pilots were given cues dictated by either the hybrid scheme or one of two competing schemes (cause or flight phase). We then assessed the degree to which each scheme promoted recall accuracy.

The procedure was the same as used in Study 3, except for during the recall phase. The 36 pilots were each randomly assigned to one of the four experimental conditions: the no cues control condition, the hybrid cues condition, the flight phase cues condition, and the causes cues condition. After the last cue was provided to pilots in the hybrid, flight phase, and causes conditions, pilots were asked if they could think of any other events.

Pilots in the flight phase condition were presented with a series of recall cues corresponding to eight phases of flight (e.g., "Can you think of any events during pre-flight?", "Can you think of any events during take-off?"; Cues: preflight, taxiing out and in, take-off, climb, cruise, descent, approach, landing). Pilots in the cause condition were presented with a series of recall cues corresponding to four potential causes of events (e.g., "Can you think of any events involving equipment malfunctions?", "Can you think of any other events caused by people or human error?"; Cues: equipment malfunctions, people, conflicts with other vehicles, weather). Pilots in the hybrid condition were presented with a series of recall cues corresponding to eight hybrid clusters (e.g., "Can you think of any events involving equipment malfunctions?", "Can you think of any events involving turbulence?"; Cues: equipment malfunctions, turbulence, weather, passengers, air traffic conflicts, events on the ground, problems caused by the cockpit crew, altitude deviations). The cues were mentioned aloud by the interviewer and were simultaneously shown to pilots on 3 x 5 index cards.

If the hybrid organization does in fact well-characterize pilots' memory organizations, then the hybrid cues should have significantly enhanced recall. To test this hypothesis, we scored the transcripts of pilots' recollections in the four experimental conditions for the number of specific events correctly recalled. To compute a total accuracy score for each pilot, we first counted the number of recollections of specific events that were clearly and unambiguously identifiable as one of the 96 events the pilot had read and if all information recalled about that event was recalled accurately, even if the pilot failed to mention some of the information contained in the written description of the event.

We then added to this the number of general categories correctly mentioned (e.g., "There were several turbulence problems") by the pilot. Such recollections represent some accuracy, because they indicate that the pilot correctly recalled some of the information presented, even though not in terms of specific events. However, we did not count a general category statement toward the pilot's total accuracy score if (a) a specific event that was an instance of the general category was mentioned following the mention of the general category, or (b) the general category was simply a restatement of a cue that had been provided to the pilot by the interviewer (e.g., a pilot saying

“Well, there were some turbulence problems” after being given the “turbulence” cue by the interviewer). Any specific event or general category that was mentioned more than once by a pilot was counted only once. Any attempt at recalling an event that included some correct information and some incorrect information was considered to be incorrect.

As expected, the hybrid condition’s mean total accuracy score was sizably larger (27.22) than the means for the other three conditions, and surprisingly, the cause condition’s mean (17.89) was lower than the means for the flight phase (21.78) and no cues (21.11) conditions. Thus, recall accuracy was 34% better in the hybrid condition than in the average of the other three conditions ($M = 20.26$). The number of events correctly recalled differed significantly between conditions, $F(3,32)=3.21$, $p=.036$. Recall in the hybrid condition was reliably greater than recall in the cause condition ($p=.054$) and the no cues condition ($p=.005$) and marginally significantly greater than the flight phase condition ($p=.085$). The no cues, flight phase, and cause conditions did not differ from each other even marginally significantly (smallest $p=.21$). Furthermore, a planned contrast indicated that recall accuracy was higher in the hybrid condition than in the other three conditions combined, $F(1,32)=7.78$, $p=.009$. All this provides strong reinforcement for the conclusion that the hybrid cues facilitated recall the most and supports the effort to empirically identify the organization scheme used by pilots rather than use a priori, though reasonable, potential organization schemes as cues in a subsequent survey.

General Discussion

These studies suggest that survey researchers should consider pretesting to identify the organizations of events in respondents’ memories when designing questionnaires, because our a priori guesses about mental organization, though well-informed, turned out to be incorrect. Armed with converging assessments of memory organization, we found that recall was enhanced when a questionnaire employed matching cues. This both reinforces psychologists’ notions of memory organization and provides a practical and useful recommendation to survey designers.

References

Arabie, P. & Carroll, J. D. (1980). MAPCLUS: A mathematical programming approach to fitting the ADCLUS model. Psychometrika, 45, 211-235.

Brewer, M. B., & Lui, L. N. (1996). Use of sorting tasks to assess cognitive structures. In N. Schwarz & S. Sudman (Eds.) Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research (pp. 373-385). San Francisco: Jossey-Bass.

Cohen, J., & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82, 407-428.

Ostrom, T. M., Carpenter, S. L., Sedikides, C., & Li, F. (1993). Differential processing of in-group and out-group information. Journal of Personality and Social Psychology, 64, 21-34.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. Psychological Bulletin, 76, 45-48.

Walker, M. E. (1998) Standal: A program to format proximities and cluster memberships for performing stand-alone regression (Version 2.0) [Computer program]. Columbus, OH: The Ohio State University.