Proceedings of CHI'96, the ACM Conference on Human Factors in Computing Systems, ACM Press, New York

The Group Elicitation Method for Participatory Design and Usability Testing

Guy A. Boy
European Institute of Cognitive Sciences and Engineering (EURISCO)
4, Avenue Edouard Belin, 31400 Toulouse, France
Tel: (33) 62.17.38.38
E-Mail: boy@onecert.fr

ABSTRACT
This short paper presents the Group Elicitation Method (GEM), a brainwriting technique augmented by a decision support system for participatory design and usability testing. GEM has been successfully used in four industrial projects to elicit knowledge from users, management and designers. In particular, in three of them it was used to elicit end-users' knowledge for the design of new user interfaces. This short paper discusses the properties of such a method and the lessons learned.

KEYWORDS: Knowledge elicitation, participatory design, decision support systems, evaluation, methodology.

In industrial environments, several experts may work together to investigate appropriate solutions to a design or usability issue. A crucial problem is to derive an acceptable consensus from a group of experts who share neither the same background nor the same objectives. It is not uncommon that experts do not understand each other. Strong personalities may dominate a meeting even if they do not contribute much from a content viewpoint. It often follows that group decision making is not always democratic. The *Group Elicitation Method* (GEM) has been designed to moderate these effects. Six to ten domain experts (users, managers and designers) are chosen to participate in the GEM experiment. Experience suggests that the optimal number of participants for an interesting session in a reasonable time frame is about seven.

A typical GEM session takes a full day, and consists of six phases conducted by a knowledge elicitation facilitator:

- issue statement formulation and choice of the participants (We found that the type of checklist proposed by Nielsen *et al*. [3] for semi-structured interviews was a good starting point for the formulation of issue statements. GEM currently uses the following checklist as a starting point: what is the goal of the engineered system that we plan to design or evaluate?; how is the system or its equivalent being used? (current practice, observed human errors); how would you use this system? (users' requirements); what do you expect will happen if the corresponding design is implemented? (e.g., productivity, aesthetics, and safety issues); how about doing the work this way! (naive and/or provocative suggestions); what are the constraints that you foresee? (pragmatic investigation of the work environment);
- viewpoints generation (A premise of this phase, better known as *brainwriting*, was introduced by Warfield [4] to facilitate the generation of ideas, and further developed by Boy [1]. This phase usually takes one hour for seven participants);
- reformulation of these viewpoints into more elaborate concepts (This phase involves four types of operations that correspond to the concept clustering mechanisms that were described and used in Fisher's COBWEB system [2]: classifying the viewpoint with respect to an existing concept (a class of viewpoints); creating a new concept; combining two concepts into a single concept; and dividing a concept into several concepts. This phase may take between one and two hours.);
- generation of relations between these concepts (Each participant is then asked to provide his opinion on the *relative priorities* of these concepts. He/she needs to fill out a triangular matrix presenting the concepts in rows and columns. Basically, each matrix "box" is filled by a score that can be *+1* / *0* / *-1* if the line concept is *more* / *equally* / *less* important than the column concept. This phase takes about thirty minutes.);
- derivation of a consensus (A *consensus* is derived using an interactive computer program. We call *global score of a relation* the sum of all the scores of one relation among the participants. The global score matrix is the sum of all the matrices generated by all participants. To each global score is attached a standard deviation measuring the inter-participant consistency of the global score of the relation. This phase leads to the expression of a consensus among the participants. This consensus is expressed using four types of typical parameters.);
- critical analysis of the results (We observed that a critical analysis of the generated concept network is very constructive and reinforces the consensus. Experience shows that this phase may take between thirty minutes and one hour. A report is finally prepared by the knowledge elicitation facilitator.).

GEM PROPERTIES AND LESSONS LEARNED

Participants involved in a GEM experiment, easily but purposefully write short statements that usually include stronger viewpoints than if they were giving the same information by talking. Writing involves loose ratification. Generated viewpoints are more or less specific episodes that can be subsequently endorsed or contradicted. As the GEM progresses in time, participants start to write for the others instead of writing for themselves by adapting their own language to what they have just read. Participants need to understand the needs and abilities of the others if they want to communicate. The difference between talking and writing lies in the fact that they have more time and opportunity to express their viewpoints in an undisturbed atmosphere.

During the viewpoint generation phase, by reading each others' statements, each participant allows his/her attention to be focused on the precise details of the argumentation. This *read and write* dialogue also involves both an event-driven reactive behavior and a reflective/deliberative behavior. In any case, the fact that people are required to write statements, read other's statements (support or refute), add new statements, and repeat until enable generate new meaningful contributions involves cognitive processes of interpretation, understanding and reformulation. Positive reinforcement contributes to the validity of the viewpoint. A negative statement or refutation provides an alternative view but there is an assumption of position by virtue of being written–the argument against must be strong enough to dislodge it. In this context, it is possible to talk about a logic of *discovery*. The indeterminate nature of the result of interactions between the participants as rational agents creates a situation of investigation and confrontation of knowledge. By reading each others' statements and, in the later phases of GEM, by visualizing conceptual structures that the group generates, participants need to re-examine their own knowledge. The resulting synthetic effect is a crucial added value of GEM: participants learn, and subsequently are more motivated to accept the result. This was observed in each GEM experiment that we carried out. In the beginning, most participants were very skeptical about the use of GEM since they thought that they could not come to an agreement. GEM not only provided a consensus, but also an indisputable written agreement. All groups were very enthusiastic after each GEM experiment.

Visual feedback enables participants to explicitly perceive the course of the knowledge design process. Participants see the viewpoint generation as a sequential process, i.e., they generate their own viewpoints and react to each other's viewpoints using the sequentiality of writing. It is also a parallel process since all participants generate viewpoints at the same time (reducing elicitation time). The difference between a regular meeting and this phase is that the participants discuss viewpoints that they have already explicitly expressed and were captured. The ontology of the domain that is explored is cognitively manipulated by reading and writing. This is extremely different from verbal expression of the same viewpoints; verbal expression is usually ephemeral.

GEM is grounded in social interactions mediated by an incrementally-constructed shared memory. During the generation phase, participants are collaboratively involved in a creative activity using the sheets that are progressively filled in. During the reformulation phase, they cluster the viewpoints into concepts that are compared during the relation generation phase. Generated concepts are *good enough* solutions that are accepted by the participants. Consensus ratings are computed and used as mediating relative trends to guide designers and evaluators in their final decision making process.

We elicited several criteria used for such a construction of knowledge (during the reformulation phase):

- *Simplicity*. For reasons of comprehensibility, if the expression of a concept is too complex, it must be split into several simpler concepts. The simplicity of a concept is usually a function of the number of viewpoints that generated it.
- *Interest*. The interest of the participants for a concept is provided by the number of positive reinforcements of its various viewpoints. It may be considered also to be inversely proportional to the number of its strongly contradicted viewpoints.
- *Robustness*. The robustness of a concept is a measure of how much the participants support it. This, in turn, depends on the robustness of the initial viewpoints and backgrounds of the participants.
- *Corroboration*. By corroboration, we mean validation by exterior knowledge which is provided during reformulation discussions. We here consider that a concept is corroborated if other concepts reinforce its constituting viewpoints.

ACKNOWLEDGEMENTS

Gene Lynch and Helen Wilson provided astute advice towards improving the quality of this paper.

REFERENCES

1. Boy, G.A. (1991). *Intelligent Assistant System*. Textbook. Published by Academic Press, London.
2. Fisher, D.H. (1987). Knowledge Acquisition via Incremental Conceptual Clustering. *Machine Learning*, 2, pp. 139-172.
3. Nielsen, J., Mack, R.B., Bergendorff, K.H. & Grischlowsky, N.L. (1986). Integrated software usage in the professional work environment: evidence from questionnaires and interviews. *CHI'86 Conference Proceedings*, Mantei M. and Oberton P. (Eds.), pp. 162-167. New York: ACM Press.
4. Warfield, J.N. (1971). *Societal Systems: Planning, Policy and Complexity*. Wiley, New York.