Pyramiding: Efficient search for rare subjects

"Pyramiding": Efficient identification of rare subjects

Eric von Hippel*, Nikolaus Franke**, and Reinhard Prügl***

MIT Sloan School of Management Working Paper 4719-08 October, 2008

ABSTRACT

The need to economically identify rare subjects within large, poorly-mapped search spaces is a frequently-encountered problem for social scientists and managers. It is notoriously difficult, for example, to identify "the best new CEO for our company," or the "best three lead users to participate in our product development project." Mass screening of entire populations or samples becomes steadily more expensive as the number of acceptable solutions within the search space becomes rarer. Pyramiding can be significantly more efficient under many conditions.

"Pyramiding" is a search process based upon the view that people with a strong interest in a topic or field tend to know people *more* expert than themselves. In this paper we empirically explore the efficiency of pyramiding searches relative to mass screening in search spaces where there is only one "best" solution. In four experiments, we find that pyramiding in each case identifies the best solution within the search space using an average of only 30% of the effort required by mass screening. Based on our findings, we propose conditions under which pyramiding will be a more efficient approach to exploring a search space than screening.

* MIT Sloan School of Management, Cambridge, MA, USA: evhippel@mit.edu

** WU Wien (Vienna University of Economics and Business Administration), Institute for Entrepreneurship and Innovation (www.e-and-i.org), Vienna User Innovation Research Initiative (www.userinnovation.at), Austria: nikolaus.franke@wu-wien.ac.at

*** Innsbruck University School of Management, Marketing Group of the Department of Strategic Management, Marketing, and Tourism, Austria: reinhard.pruegl@uibk.ac.at

"Pyramiding": Efficient identification of rare subjects

1 Introduction and overview

Identification of subjects with rare attributes within large, poorly-mapped search spaces is a frequently-encountered task in social science research. Mass screening, a common search approach, involves collecting information from *every* member of a population or sample to identify those with the desired attributes. Clearly, as individuals with the desired attributes become rarer in a population, the number of people who must be screened to attain each "hit" increases, and screening becomes an increasingly inefficient mode of data collection. As Sudman puts it: "If the [desired] population is rare or very rare, screening costs may be very large and account for the major share of data collection costs" (1985, p. 20). Under such conditions, a more efficient method is needed.

One method which has been found useful to efficiently identify people who have a rare attribute in common is "snowball sampling" (Goodman, 1961). Snowball sampling involves asking individuals who have a rare characteristic being sought to identify others they may know who have that same characteristic (Welch 1975). The utility of snowballing stems from the observation that people tend to know or be aware of people like themselves.

Pyramiding is a variant of snowballing – but with an important difference. Pyramiding is a search process based upon the idea that people having a strong interest in a given attribute or quality, for example a particular type of expertise, will tend to know of people who *know more about and/or have more of that attribute* than they themselves do (von Hippel et al 1999). For example, if an individual is an expert heart surgeon, pyramiding assumes that that individual will know of others who are still more expert in that field. Similarly, if a person is an avid collector of jazz CDs, pyramiding assumes that person is likely to be able to identify people with still larger collections of jazz CDs.

Pyramiding is useful when a researcher wants to efficiently identify the persons with higher *levels* of a given attribute in a population or sample – generally individuals near or at "the top of the pyramid" with respect to that attribute. The pyramiding process is quite simple in concept: one simply asks an individual to identify one or more others who she thinks has higher levels than she does of the sought-after attribute – or better information regarding who such people might be. The researcher then asks the same question of the persons so identified, and continues the process until individuals with the desired high levels of the attribute have been identified.

As an example of where pyramiding has proven its usefulness, consider market researchers' problem of identifying "lead users" within a population of product users who can help manufacturers to develop new products. Lead users are relatively rare in a population, and are defined as having high levels of two attributes relative to the population average: They are (1) at the leading edge of an important market trend(s) and (2) they have a high need for solutions to the novel needs they have encountered at that leading edge. Early studies seeking lead users used a screening method. However, the high cost of that methodology quickly became evident. Thus, Lüthje (2000) used screening methods in lead user studies in two areas (game development and public transportation). Altogether, he contacted 2,043 persons and, based on carefully developed quantitative criteria, identified 22 lead users among them as suitable for further detailed analysis. This means that the sampling efficiency was only 1.1% - 98.9% of all contacts were "waste." As a result of these low efficiencies, those conducting lead user studies have increasingly turned to the pyramiding method for lead user identification (e.g. von Hippel et al., 1999; Olson and Bakke, 2001; Lilien et al., 2002).

Even though pyramiding is increasingly used in the field, its efficiency relative to mass screening has never been measured or otherwise explored. Clearly, it is important to do this if pyramiding is to become an important part of researchers' toolkits – and so in this paper we begin that work. We proceed as follows. In section 2, we further explain pyramiding and screening and give an illustrative example of the two methods. In section 3 we review related literature. In section 4 we report upon study that allowed us to simulate 663 "real" pyramiding search chains in 18 search settings and compare the efficiency of these with screening methods applied to the same settings. In section 5 we discuss our findings, and derive propositions about under what circumstances pyramiding is a more efficient search method than screening.

2 Pyramiding vs. Screening

2.1 Background on pyramiding and screening

Pyramiding, as was mentioned earlier, is a variant of "snowball sampling" (Goodman, 1961,Welch 1975). Snowballing assumes that people in any population will tend to know others similar to themselves. The snowballing method therefore begins with a few people in a population known to have a given rare attribute, and asks these people for the identity of others that have that same rare attribute. Pyramiding, unlike snowballing, enables searchers to move "up the pyramid" – to find people with *more* of a given attribute – rather than staying at the same level (von Hippel et al 1999).

Applying a screening questionnaire to a group of n people is in essence an experimental protocol involving n experiments conducted in parallel. Pyramiding, on the other hand, is an experimental protocol involving x experiments conducted in series. In the case of experiments conducted in parallel, no learning is possible between experiments. In the case of experiments conducted in series, it is possible for a researcher to incorporate learning acquired from previous experiments into each succeeding experiment in the series.

“Hill-climbing” is a form of serial search for a solution where learning from each experiment is incorporated into the next in the series (Thomke et al., 1998). In a standard hill-climbing method, an experimenter moves across a landscape in which desired solutions can be found at the tops of "hills" on that landscape: the higher the hill, the better the solution found at the top. The experimenter takes one step at a time, with each step representing an experiment. After each step, the experimenter learns from that experiment by determining which of his feet is at a higher point on the landscape. This learning is incorporated into the next experiment in that the experimenter turns towards the higher foot before taking the next step. A hill-climbing strategy enables an experimenter to travel to the highest point that can be reached by a continuously ascending path in the topography that he encounters in his travels. However, one major disadvantage of the strategy is that she will not know whether the highest point reached is in fact the highest peak on the landscape or simply a foothill.

“Pyramiding” involves the same basic hill-climbing metaphor as just described (figure 1). However, each location on the hill reached by the experimenter is not just a physical point in the landscape, but an intelligent person who has some level of knowledge about the surrounding terrain. After taking a step, the experimenter tells the person encountered at that spot about her desired goal. That person may be able to respond with very useful information about the terrain, such as: ”I know that X innovation or innovator is at the top of this hill – jump directly there.” Or, “Now that you have told me about the type of solution you are seeking, I can tell you that you are not on the best hill – there is a higher hill over there.”

In many common market research applications, protocols involving parallel experimentation are appropriate because learning between trials is not useful and only costs time. For example, parallel experimentation is appropriate to determine the distribution and intensity of average characteristics or information in a specified group – typically a target market. “We want to know how many people in this target market are aware of our brand, how high their brand attachment is, and how they would react to a brand extension.” To answer such questions, there is no benefit from learning what individual A thinks prior to collecting information from individual B. In contrast, when searching for persons with rare levels of high expertise in a population – our case – a serial experimental protocol may be more appropriate than a parallel one. In this case applying what is learned in one "trial" to the next may be very useful because individual A may well have information that can guide researchers very quickly to the rare "top expert" being sought.

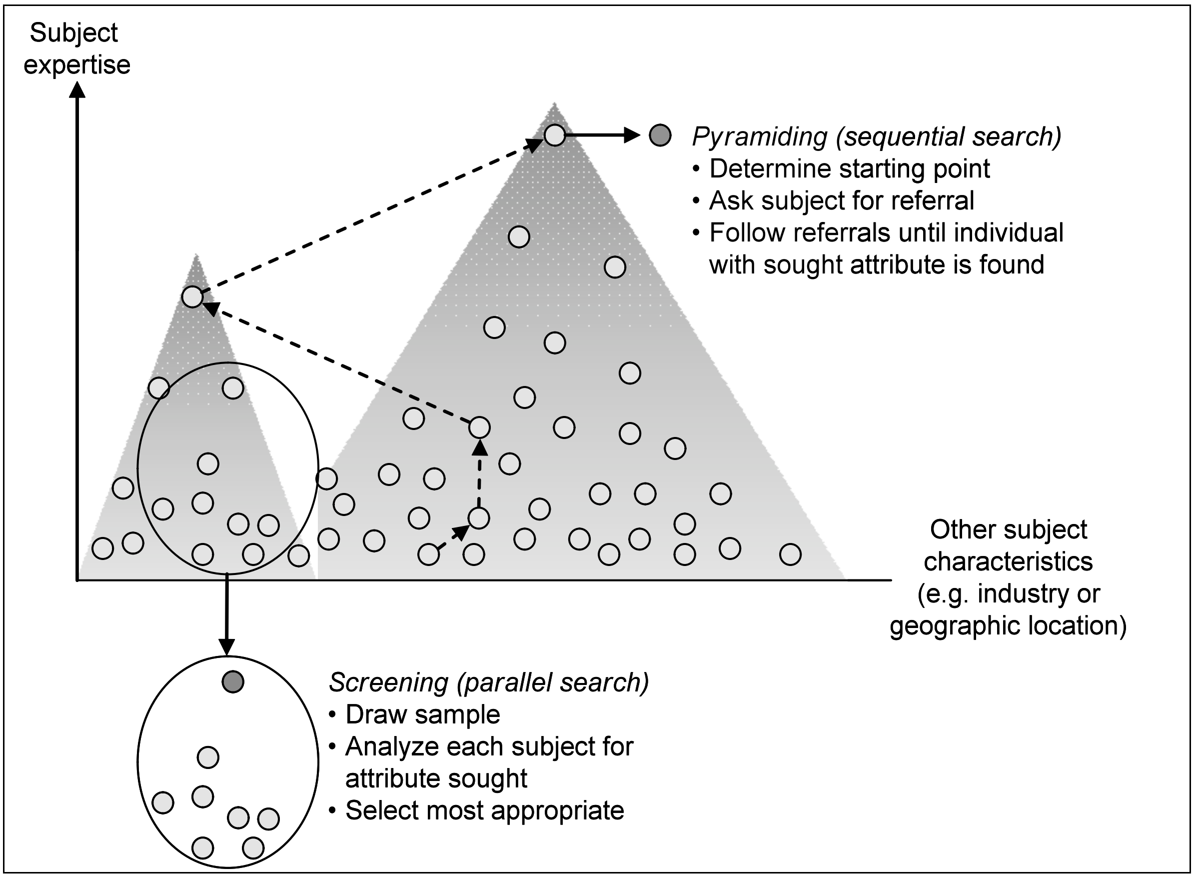

Figure 1: The search concepts of screening and pyramiding

2.2 An Illustrative Case Study of Pyramiding vs. Screening

IDC is a firm specializing in collecting and disseminating information about information companies and markets. A project carried out at IDC was used by John Gantz, a senior manager at IDC, and one of the authors as an opportunity to compare the effectiveness of screening and pyramiding techniques in the identification of lead users. The goal of the IDC project was to identify users at the leading edge of six trends previously identified by IDC as important to the evolution of Internet and Intranet websites. Approximately 12 lead users were required, two at the leading edge of each of the six trends. Once identified, these 12 individuals were to be invited to a lead user workshop sponsored by IDC to discuss new technical advances occurring or likely to occur in the website field.

Identification of these lead users promised to be a difficult task. No public data was available on corporate website performance parameters. Nor was there any data available on the presence or absence of innovative work by those managing corporate sites. In addition, of course, individuals with appropriate lead user and presenter characteristics were likely to be rare. Compounding the difficulty, a date for the workshop had been set and publicized which allowed very little time to identify and recruit the lead users needed as participants. Given this situation, John Gantz decided it would be prudent to conduct parallel recruitment efforts – one using IDC's standard screener methodology, and one using the pyramiding method. Both recruitment efforts were managed and partly staffed by a skilled project team of four IDC professionals.

For the purposes of this project, lead users were defined as individuals who (1) were directly involved with a corporate website that was at the leading edge of one or more of the six important trends in the website field (such as the "hit rate" a website was experiencing) and (2) had a strong incentive to improve website performance in that dimension, as evidenced by their personal record of innovation (such as writing novel computer code to improve the performance of their site in that dimension). An additional requirement placed on the lead users who would ultimately be selected as workshop participants was that those recruited should be "likely to be interesting presenters and discussion partners" – a matter to be assessed by telephone discussions between an IDC project team member and each potential recruit.

Screener method. The total US population of webmasters was probably somewhere around 100,000 at that time. The starting population for the IDC screening procedure was 2,000 names of webmasters assembled from three sources. Two were published lists of webmasters, and the third was a list of 100 names of especially well-qualified webmasters that had been compiled by a project team member during a prior research project on the Internet. A screening questionnaire was created by the IDC project team to collect preliminary information as to whether a webmaster contact was possibly a lead user. This questionnaire was to be administered by employees of an outside market research firm during 15-minute telephone interviews with the individuals on the list. Respondents who answered in ways that indicated they might be lead users were then to be contacted and interviewed further by an IDC project team member so that a final determination could be made.

Four hundred 15-minute screener interviews were in fact carried out by the outside market research firm before the screener process was stopped due to the greater success of the pyramiding method. The 400 screener interviews yielded 25 likely lead users who were interviewed further by IDC project team members. Three of these 25 interviewees were found to be actual lead users and were selected for participation in the IDC Lead User workshop. (One of the three was also independently identified by the pyramiding method.) John Gantz computed the costs to IDC of the screener-based recruitment effort from its start to its early termination and came to a figure of $29,300. The cost of the screener method per lead user identified was thus $9,767.

Pyramiding method. The starting population for the pyramiding method was the initial list of 100 selected webmasters plus 50 to 100 appropriate personal contacts known to IDC project team members. The initial IDC project team review and discussion of these potential interviewees resulted in the selection of two dozen or so as useful starting points for the pyramiding method. Pyramiding interviews were then carried out by IDC team members. Approximately 80 to 100 phone interviews were conducted, and eight lead users were found who were qualified for participation in the IDC lead user workshop. Two of these were among the initial starting points selected by the team for the pyramiding process; two were second node contacts (Fedex, Citibank); and four were third node contacts (Sandia Labs, Virtual Vineyards, Time Warner, US West). The cost of the pyramiding process to IDC was calculated by Gantz to be $12,000 in IDC personnel time. Thus, the cost per lead user identified via the pyramiding procedure in this real world case was $1,500 – 15% of the cost of the screener method.

3 Literature review

To the best of our knowledge, there is no prior work that compares the efficiency of pyramiding search methods relative to screening search methods. However, useful related work does exist, and we review this next.

Pyramiding involves traversing reputation networks, and the goal is to do this in an efficient way. The field of scholarship that sheds light on this type of endeavor is called network theory or graph theory (for a recent review: e.g. Newman 2003). Networks consist of nodes (in our case a sample of people) and links between the nodes (in our case reputational information held at each node about other nodes). Network theory points out that the efficiency with which one can move from point A to desired point B will depend upon the structure of the network – the number and distribution of the links between nodes (e.g. Burt 1992, Strogatz 2001). For example, consider a "star" network. In this network structure, one node will be at the center and all other nodes will be linked to each other only via the central node – there are no direct links between peripheral members. In the case of a star network, one can immediately see that the number of links required to reach a desired peripheral member from any other peripheral member will always be two: a first step from any peripheral member to the central member who has information on all peripheral members, and then a second step to the desired member. In contrast consider a "chain" network in which all nodes are connected by a single chain of links. In that case, clearly, it will take as many steps as there are intervening nodes to move from a starting point to a desired end point on the chain.

The most efficient pathways through "pure types" of networks like a star network can be calculated. However, the most efficient pathway through real-world networks is much less predictable, because the shortest path from one point in a network to another can be drastically affected by the addition of even a single additional link. For example, one link added to directly connect two distant points in a chain network can clearly have a great impact on many "shortest paths" in that network. For this reason, experiments with real-world networks tend to require "try it and see" to determine the properties of a given network.

With respect to informant identification, the method most closely related to pyramiding is, as we mentioned earlier, snowball sampling or "snowballing" (Goodman, 1961). Snowballing is occasionally used in marketing and other fields to assemble samples of individuals with characteristics which are relatively rare or hidden in a population (Spreen 1992). Among these are special populations such as the deprived, the socially stigmatized and the elite (Atkinson and Flint 2001). Thus, Burt (1982) used snowballing to identify experts in a specialized academic field in *addition* to those who had explicitly labeled themselves as belonging to that field and who had high reputation on a range of measures. These "non-labeled" people were identified by "labeled" experts as fellow experts worthy of inclusion in their group. Snowballing is seen in market research as way to obtain samples of special populations in an efficient manner (Sudman 1985).

Snowballing capitalizes on the reputation networks of sample participants. Its efficiency depends upon the common observation that people tend to know or be aware of people who are like themselves. Sample identification proceeds by first finding one or a few individuals who have the rare characteristic. Then one asks these individuals to identify others they know who have that same characteristic (Welch 1975). In this way, the researcher creates a growing pool of contacts (Atkinson and Flint 2001). Researchers using snowballing must actively develop and control the initiation, progress and termination of the sample (Biernacki and Waldorf, 1981). Welch (1975) provides some empirical evidence on the relative efficiency of snowballing versus screening. The author tried to locate Mexican-American households in Omaha, Nebraska. This target group was thinly dispersed (8,000 households out of 400,000). A combination of screening and snowballing was used (households were randomly chosen until a Mexican-American household was found; the interviewed person was then asked for referrals). Over one third of the respondents provided referrals, with an average of two each (Welch 1975). Although only 54% of all households were screened, 77% of the targeted households were located through this combination of screening and snowballing. A comparison of the households located through screening (n=87) and snowballing (n=61) showed no significant differences in sample characteristics. The method used in this study was able to reduce the number of unproductive calls made by interviewers and therefore lowered the costs involved.

Another phenomenon related to pyramiding – but one that differs in crucial respects – is the "small world" procedure first proposed by Stanley Milgram in 1967. This procedure has been used to study acquaintanceship networks in a population. In small world studies, the underlying idea is that one can reach any person in a population (with the exception of totally isolated individuals or subgroups) by passing a message along acquaintanceship chains. Thus, in 1969, Travers and Milgram explored how many acquaintanceship links it would take to transmit a message from an arbitrary starting point to a known end point in a population. They defined an acquaintance as someone "personally known to you on a first-name basis". They then asked study participants (for example, some individuals living in Nebraska) to try to transmit a letter to a specified "target" person at a specified address (which was in Boston) by mailing or delivering it to an acquaintance who they thought was most likely to know the target person. That message recipient was to take the same actions in turn, and the chain was to continue until the letter had been delivered to the target person via this "acquaintanceship chain". Travers and Milgram (1969) found that 29% (64 out of 217 started) of the chains reached the target person, with successful chains requiring an average of approximately 5 acquaintanceship links to reach the goal.

Dodds, Muhamad and Watts (2003) conducted a large internet-based social search experiment in the Travers and Milgram style. Using 18 target persons from 13 countries, the authors recorded data on 61,168 individuals from 166 countries generating 24,163 message chains. They found a median of 5 to 7 steps to reach the target person. On the other hand, the completion rate was rather low – only 384 out of 24,163 chains reached the target (1.6%). Successful searches were enabled by weak ties and relied to a great extent on professional relationships. Those who forwarded the message were found to have chosen the next link in the message chain mostly on two criteria: geographical proximity to the target and similarity of occupation.

Small world networking studies differ from pyramiding studies in two important respects. First, messages in small world studies are intended to flow along *acquaintanceship* chains, where acquaintances are defined as personal relationships. In pyramiding studies, in contrast, participants in chains leading to the identification of experts are *not* required to be personally acquainted with the next link in a chain that they identify. Instead, it is useful to link also to those whom they only know by reputation. Abandoning the requirement for actual acquaintanceship can greatly increase the efficiency with which desired subjects can be identified via pyramiding. For example, essentially everyone in the heart transplant field can immediately identify the top practitioners in that field, even if only a few are personally acquainted with them. Kleinberg (2000) shows mathematically that individuals are able to find short chains in a large network based on local information.

Second, in small world studies each link in the chain must have the *motivation* as well as the information needed to forward a message along to the next link. In pyramiding in contrast, contacts with each chain member are made by a researcher who has an independent motivation to pursue the chain to the end. The low chain completion rates found in small world studies are therefore not likely to be a problem in the case of pyramiding studies. A study by Guiot (1976) illustrates. He conducted a small world study in which he used a telephone procedure instead of mailing. Pursuing each identified link in a chain was entirely at the initiative of and under the control of the researcher, a situation very similar to that in pyramiding studies. A small pilot study (52 starting persons) yielded a completion rate of 85%. Dodds, Muhamad and Watts (2003) stress the importance of incentives for individuals to increase search success.

4 Research methods

Recall that our overall goal is to empirically compare the relative efficiency of pyramiding and screening methods with respect to identifying the single individual in a group who has the "most" of a given attribute. In outline, our research method involved selecting relatively small groups of individuals for study. Within each of these populations we collected two types of information from *every* person in the group: (1) the individual's actual level of the attribute being asked about and; (2) the identity of another person within the group which each subject thought had the highest level of that attribute of anyone in the group. For example, in one of our studies, the attribute of interest was "Who in this group has the most jazz CDs?" Each individual in that group was therefore asked: (1) "How may jazz CDs do *you actually have*?" and (2); "Who in this group do *you* think has the most jazz CDs?"

Having obtained these two bits of information from each individual, we could then simulate all pyramiding chains within each group, and determine the efficiency (number of nodes from start to end point) of each simulated search. We could also determine whether each search was successful – that is, which of all the pyramiding chains in the group ultimately identified the target individual – the one who actually had the most of the attribute at issue. Screening costs were then very simply determined from the total size of the group, and efficiency of pyramiding vs. screening search strategies could then be determined.

4.1 Populations

We included four different groups – independent populations - in this study. All groups were relatively small, and we were able to obtain data from *all* members of each group. (This was important: any person unavailable to or refusing to answer would potentially ruin the simulation of many chains.) Groups were intended to be different in nature in order to allow for some variance in our findings. A total of 147 individuals participated in our study.

	Population 1: Chorus	Population 2: Football team	Population 3: Teaching staff	Population 4: Students' association
Meeting frequency	weekly	2-3 times a week	daily	monthly
Meeting purpose	hobby	hobby	work	socializing
Age	12-18	20-30	30-60	20-35
Size	41	33	38	35
Male	12	33	29	35
Female	29	-	9	-

Table 1: Populations studied

4.2 Search Topics

We used in total six different attributes or "search topics" in our studies (table 2).

Topic	Individual searched for
(1) Mountaineering	The individual in the group who had climbed the highest mountain (measured

	in meters)
(2) Jazz	The individual in the group who has most jazz CDs (measured in numbers)
(3) Weakness of vision	The individual in the group who has weakest eyes (measured in diopters)
(4) Car accident	The individual in the group who had had the most car accidents (measured in number of car accidents within the last 5 years)
(5) Stay in hospital	The individual in the group who had the longest stay in hospital in his or her life (measured in month and/or weeks and/or days)
(6) Apartment size	The individual in the group who has the biggest apartment (measured in square meters)

Table 2: Search topics used in studies

Due to constraints on group members' time, members of populations 1 and 2 one participated all 6 topics; members of populations 3 and 4 participated in only 3 topics each. Taken together, responses from group members in the 18 topic studies enabled us to simulate a total of 663 search chains for study (41 members of population one*6 topics + 33 members of population two*6 topics + 38 members of population three*3 topics + 35 members of population four*3 topics).

4.3 Data Collection Procedure

Data collection was done by means of written questionnaires. Each subject was asked: (1) about their own information regarding the topic at issue; and (2) whom they would refer us to: "We are looking for the one subject in the group [who has climbed the highest mountain]. Please refer us to the group member who you think might be the person we are looking for, or to a group member who you think has better information on this topic than you do." Subjects could only refer to persons within the group (who were recorded on a name list). The comprehensibility of the questions was checked using pre-tests.

During the experiment we carefully controlled for two very important possible sources of bias: First, recall that we ensured that all group members were present when we asked each group to fill out the questionnaire. This was necessary because, if a key member referred to by others was missing, many pyramiding chain analyses would be impossible to complete. Second, we ensured that each group member filled out the questionnaire simultaneously and completely independently. Independence of referrals was, of course, a necessary precondition for our simulation.

4.4 Measurement

Efficiency of a pyramiding search. We define the "efficiency" with which the person holding the greatest amount of an attribute is identified in a population as the number of contacts that must be made before that person is correctly identified. In the case of a parallel screening search, the number of contacts that must be made will equal P where P is the number of individuals in the population. In the case of pyramiding one ends the search if a contact identified is the target person – that is, has the highest level in the group of the attribute sought for. This can best be explained using an example. Figure 2 shows a typical system of referral patterns. Subject #20 is the one person whom we had identified as the person who climbed the highest mountain. When asked, this person had correctly referred to herself as the searched-for person.

Referral chains starting with subjects #1, #4, #13, #14, #15, #16, #17, #18, #22, and #35 would be classified as "unsuccessful" because they ended with the wrong person (#23). The search chain would terminate with the wrong person in such cases because she (or he) incorrectly identified herself as the person being sought for – and so would not refer searchers on to a next link in the chain. This corresponds to a hill-climbing search in which only a local peak was reached. As the level of expertise is below the desired level, we would continue with another starting point.

All the other potential chains reached the target person successfully. If we e.g. had asked subject #25, she would have referred us to #12. #12 would have referred us to #34. #34 would have correctly identified #20 as target person. This chain has a chain length of 4, as 4 persons have to be asked in order to reach the target.

If we encountered an endless circle, it was broken: As soon as a person occurred repeatedly in one chain (which would create an endless circle), the chain was ended and the number of steps was counted, and the referral chain would be counted as unsuccessful.

If the real target person referred to someone other than himself (meaning she was unaware of her status), there were two different cases: (1) She referred to someone already in the chain. This would create an endless circle in which no new links would be added; (2) The target person referred to someone else who is not already part of the chain (with a lower level of expertise by definition). In this case, researchers would say "Ah, we seem to have reached the peak earlier. Let's stop here." The chain is successful, but one step is added (the "let's see if it still improves" step). For example, had person #20 not referred to herself but to #12, the chain length of the starting point #25 would still be 4. Had person #20 referred to #1, the chain length of the starting point #25 would be 5.

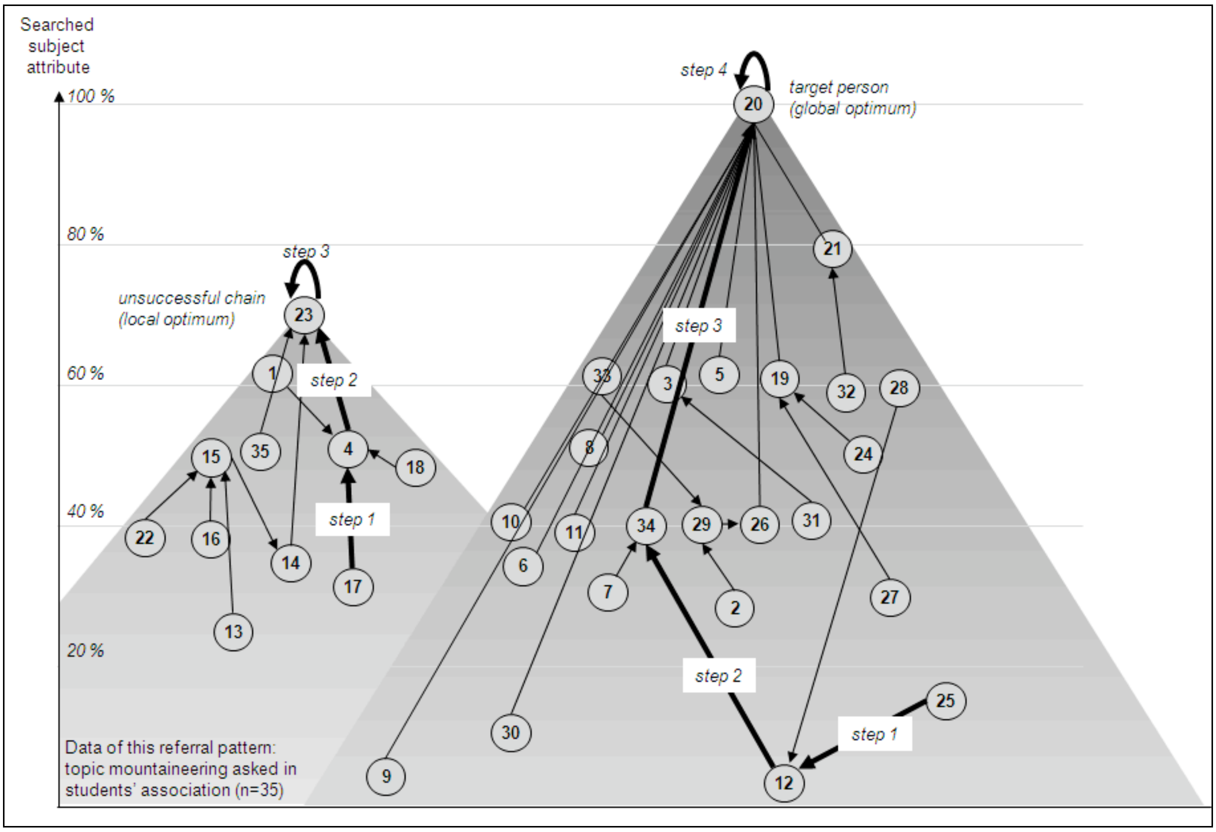

Figure 2: Example of one referral pattern (students' association; mountaineering)

Reputation of the target person. We wanted to determine whether the general popularity of the target person affected pyramiding efficiency. This popularity is termed "reputation" or "prestige" of an actor in a network in sociometric measurement literature (Wasserman and Faust, 1994). Sociometric techniques suggest providing each respondent with a fixed contact roster and asking her to describe her relationship with every individual on the roster. The reputation of a subject increases if this actor is the receiver of many connections, meaning that many individuals know this person. In addition to the existence of a directed link, we also measure "tie strength" by using a scale adapted from Reagans and McEvily (2003) (4 = I know this person very well, 0 = I do not know this person at all). Tie strength increases the likelihood of knowing more about the areas of expertise an individual has.

Reputation of a subject was then measured as follows:

(1) $$R(n_i) = \frac{\sum_{j \neq i} y_{ij}}{(g-1) * q_{\max}}$$

with:

$R(n_i)$ = standardized reputation of actor i

y_{ij} = Number of actors j connected to actor i with tie strength q_{ji}

g = total size of social network (total number of actors in a social network)

$q_{\max}$ = maximum tie strength between two actors.

Subject matter interest of population. It is plausible that referral quality will increase along with a respondent's own level of interest in the topic being asked about. This is because respondents will have higher incentives to observe and store information regarding topics that are of greater personal interest (e.g. Lavin 1965; Renninger, Ewen and Lasher 2002; Krapp 2002; Hidi and Renninger 2006). Subject matter interest was measured on the individual level by a three-item scale with a 5-point Likert-type response format (1=fully agree, 5=fully disagree). The items were "I personally find the topic of [mountain climbing] very interesting", "I really like to be involved in conversations about [mountain climbing]", and "If somebody tells me something about [mountain climbing], then I can remember this information easily." Cronbach's alpha is 0.88. Exploratory and confirmatory factor analysis showed highly satisfactory values. The individual levels were aggregated into the variable "subject matter interest of population" by calculating the mean value of individual subject matter interest with regard to each topic.

4.5 Simulation of Pyramiding Search

The expected number of contacts needed when searching for a target subject via pyramiding cannot be obtained analytically. Therefore, in order to calculate this number, we resorted to a Monte Carlo simulation programmed in C++ of a real search. We randomly selected a starting subject, X_1. If X_1 was already the target person, the number of subjects necessary to ask was 1. If not, we followed her referral to subject X_2. If X_2 was the target person, the number of persons necessary to ask was 2, and so on. If a chain broke or reached an end which was not the target person, another starting subject (other than X_1 or any of the subjects in the referral chain starting with X_1) was randomly selected and the procedure went on. The process did not end until the target subject was reached. In each system (i.e. each group and topic), 100,000 simulations of searches were conducted (total: 1,800,000). The calculation of steps necessary followed the rationale described in section 4.4.

5 Findings

To analyze the difference in efficiency between the screening and pyramiding, we compare the simulated expected number of contacts to be made when using pyramiding with the effort of screening which is by definition equal to the number of subjects in the population.

We find that the effort of pyramiding *is on average only 30%* of the effort of screening, a large efficiency gain. In all 18 cases, pyramiding involves considerably less effort than screening (table 3).

Group	**Topic**						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**Total**
Chorus (n=41)	4.16 [1.56] (10%)	26.25 [11.16] (64%)	24.03 [10.97] (59%)	16.19 [8.14] (39%)	6.17 [2.67] (15%)	14.47 [6.46] (35%)	15.22 (37%)
Football team (n=33)	5.25 [2.19] (16%)	4.49 [2.60] (14%)	9.83 [5.83] (30%)	2.95 [1.49] (9%)	21.40 [9.10] (65%)	14.50 [6.94] (44%)	9.74 (29%)
Teachers (n=38)	2.00 [0.23] (5%)	16.23 [8.61] (43%)	3.49 [1.83] (9%)	-*	-*	-*	7.24 (19%)
Students' association (n=35)	3.64 [1.93] (10%)	7.11 [3.46] (20%)	15.13 [7.45] (43%)	-*	-*	-*	8.63 (25%)
Total	3.76 (10%)	13.52 (37%)	13.12 (36%)	9.57 (26%)	13.79 (37%)	14.59 (39%)	11.27 (30%)

Pyramiding efficiency (expected value [standard deviation] of persons asked in order to identify target person, result of Monte Carlo simulation), in parentheses relative to screening efficiency.

Topics: (1) = Mountaineering, (2) = Jazz, (3) = Weakness of vision, (4) = Car accident, (5) = Stay in hospital, (6) = Apartment size

** Topic not used in this group*

Table 3: Efficiency of pyramiding relative to screening

It is very important to point out that, in real-world settings, the relative efficiency of pyramiding over screening will generally be much *greater* than we have found here. That is because, for most purposes, scholars and practitioners are seeking a person who has a "very good" level of an attribute being sought for: they are seldom seeking the person who has the *most* of that attribute relative to anyone else in a population. For example, in real world settings, one is generally seeking "a really good" new CEO, or a really good lead user, or a person with "many more" jazz CDs than the population norm. Identifying an individual meeting that kind of success criterion is generally a much easier task than meeting the search success criterion established for the studies in this paper – because there may be many "good enough" target individuals, but only one best one.

To clarify this point, recall that, in the studies we reported upon here, we surveyed all members of each of our four groups. In other words we did a full screening of these groups, followed by a simulation of all possible pyramiding chains. The findings from the screening data were what enabled us to identify the person who actually had the most of each of the attributes we were looking for our study – for example, the person with the most jazz CDs. This information, in turn, allowed us to *know* the attribute level held by the target person we would be looking for in our simulated pyramiding studies, and to judge the success of each pyramiding chain on this basis.

Of course, in the real world, pyramiding studies are *substitutes* for screening studies – and so researchers using pyramiding would not know the amount of the sought-for attribute that the "top" person in the population actually possessed. Without knowing that amount, the researcher must set a success criterion for a pyramiding study either prior to initiating it, or during the study based upon assessments of the adequacy of what has been found to that point. (An example of the latter case: "At this point in our search we have found CEO candidate M. This person meets our needs very well – so we will stop searching for still better candidates.")

There are now three possibilities: the researcher will set the success criterion at a level so high that no one in the population can meet it; at a level that only one person in the population can meet; or at a level several can meet. If (1) the attribute threshold is set higher than anyone in the population being searched can satisfy, the result will be failure for both screening and pyramiding approaches. If (2) the success level – the level of attribute sought at which you stop a pyramiding search – is equal to the level of attribute held by the *single best* person in a population, then the results in the studies we reported upon in this paper reflect the outcome – pyramiding is likely to be significantly more efficient than screening under many conditions.

Finally, if (3), the success level is set at a level of attribute that *two or more* individuals in a population possess, then the efficiency of pyramiding relative to screening will increase. To understand why this is so, it will be useful to look again at figure 2. Consider the effect of setting the attribute level for a successful search at 60% or more of population maximum instead of at 100%. In that case, the discovery of individual #23 or any of several other individuals, would be considered a success rather than a failure – and the pyramiding study could then have been terminated as a success at an earlier stage.

On a second matter, note from Table 3 that pyramiding efficiency relative to screening varied significantly among studies. In our research, the highest relative efficiency was found in the case when teachers were asked who among them had climbed the highest mountain (expected effort only 5% of screening). In contrast, when football team members were asked who among them had the longest stay in hospital, the expected effort was 65% of screening.

The fluctuation observed suggests that the relative efficiency of pyramiding will be contingent upon the presence of from one or many specific conditions or factors. As was discussed earlier, we tested two such contingent factors in this study – subjects' reputation and subjects' personal interest in the topic being asked about – and found pyramiding efficiency significantly affected by both.

With respect to reputation, recall from section 4 that, if a target person has a high reputation, i.e. many people in the group "know him or her well," the likelihood increases that more people in the group will be able to correctly identify that person as the target person sought for. We measured the effect of this factor by calculating the reputation of the target person for each of the 18 cells and correlating the resulting value with the pyramiding efficiency. We indeed find a relatively high correlation of $r = 0.371$ ($p<0.1$) between the target person's reputation and pyramiding efficiency. This suggests that it will be easier to identify people with rare expertise if many people know them well. In contrast, of course, when no one in a group knows much about the target person, it is reasonable that the efficiency of pyramiding will decrease relative to screening. Indeed, at the limit when no one knows anything about any other group member – it is a sample of strangers with no information on each other – then pyramiding picks would be random and the expected value for the relative efficiency in comparison to screening would be 50%.

With respect to the impact of an individual's personal interest in a subject, recall from section 4 that it is plausible that subjects' referrals will be more accurate if they are interested in the topic themselves. After all, a subject with a high interest in a topic is likely to be more motivated to observe and store information regarding the topic at question than a subject with low interest. In order to quantify the strength of this relationship, we correlated the average level of subject matter interest in each cell with pyramiding efficiency. Again, we find a high correlation of $r = 0.444$ ($p<0.05$), indicating that subject matter interest indeed impacts the efficiency of pyramiding search.

6 Discussion

Prior to this paper, individual researchers seeking to identify rare individuals in a population have adopted pyramiding in place of screening as a method of searching for rare subjects simply because screening was too costly – and pyramiding "seemed to make sense." In this paper we have provided a first empirical study of pyramiding versus screening, and have documented that it indeed can "make sense" with respect to economizing upon the number of contacts that must be made to identify a target individual possessing high – or even the highest - levels of a given attribute relative to the norm in a population. These findings support the more widespread use of pyramiding by researchers.

We have also seen that the relative advantage pyramiding will hold over screening will vary depending upon the knowledge and interest that group members have about one another and the attribute being sought after. A more general way to think about this latter point is as follows. In essence, the efficiency of *screening* depends only upon knowledge that one has about oneself ("Do *you* have the characteristics we are looking for?"). Logically, therefore, it should work whenever the individuals involved do know and are willing to report the requested information. (For example, screening should work when the question is, "Do you wear glasses?" but will not work if the individual does not know the requested information about himself: e.g., many will not be able to answer the question, "What is your bone density?")

In contrast, the effectiveness of pyramiding depends upon the level of knowledge that individuals have *about each other*. Therefore, it should logically vary according to how many know the target individual, whether the sought-after bit of information is observable, interesting, or for some other reason widely known, and so on. Our data on the variation in the effectiveness of pyramiding as a function of the reputation of the target person and the interest that the observer has in the sought-after information illustrates this general point. Researchers who contemplate using pyramiding should consider the likely relevance of such factors in the samples they plan to study.

6.1 Suggestions for further research

Documenting the efficiency of pyramiding serves a useful purpose in validating it as a useful research technique that can offer significant value under many conditions. Indeed, the practice of pyramiding extends far beyond formal research. For example, when journalists "network" to find the right sources for a newspaper article they are working upon, they are informally engaged in pyramiding. Many of us also engage in pyramiding in our daily lives, as when we attempt to find the best school for our child or the best doctor for our medical condition. Clearly, many practitioners as well as researchers could eventually benefit from more and deeper academic study of the topic. In what follows we suggest three general topics we think worthy of exploration.

First, further research should be conducted on the efficiency of pyramiding under a range of conditions. There are doubtless many contingent variables in addition to the reputation and interest variables we examined in this initial study. Some additional variables to consider:

- *Observability*. The efficiency of pyramiding should increase as an attribute being sought becomes more easily observable. For example, more people are likely to know who within a group wears the largest glasses than who within the group has the rarest blood type – simply because the former is more easily observed.
- *Incentive to advertise an attribute*. The efficiency of pyramiding should increase along with the incentive of a target person to "advertise" the level of a target attribute that he or she possesses. For example, a lawyer might well want to make his or her legal speciality and skills widely known in order to attract new clients. Information on socially undesireable attributes, on the other hand, are less likely to be voluntarily disseminated by a person holding them, and so for these it is reasonable that pyramiding will, other things equal, work less efficiently.
- *The existence of data bases and scoring schemes.* The efficiency of pyramiding should increase along with the number and quality of data bases and scoring schemes related to the attribute being sought via a pyramiding study. For example, more observers are likely to know the professional attributes of an academic in a given field because many data bases report on related matters: the topics of academic research papers, the quality of journals in which academics publish, and the number of citations their publications receive. Similarly, more fans of baseball and other popular sports are likely to know the detailed capabilities and performances of professional players because rich online databases exist that present the information in terms of widely know categories, such as "number of times at bat," and "number of home runs."

A second very important characteristic of pyramiding is the opportunity to learn from each node contacted and to use this learning to improve the research during the course of data collection. For example in our own research, we often apply pyramiding by telephone interview rather than via questionnaire. We do this because we are often seeking *to improve* *the question we are asking* via pyramiding rather than simply find the answer to a pre-established question such as "who has the most jazz CDs."

Thus, von Hippel et al (1999) used pyramiding to search out the most knowledgeable experts with respect to the causes of infections resulting from surgery. During the pyramiding process in that study, expert surgeons and others with international reputations in the field were interviewed. Each was asked "Who do you know who knows something more (or different) about this topic than you do?" But in addition, each was told about the goal of the study and then asked, "in your opinion, are we asking you the questions that will best get us to our goal?" What was found in that study was that, at many nodes, deep conversations with the experts had the effect of changing and refining the *questions* motivating the study in addition to identifying individuals with still higher levels of expertise. For example, one expert who had a major impact on the course of the entire study said, in effect, "You are still thinking about the problem in the conventional way. What you need to do is think about it in *this* way instead. To pursue the pathway I suggest, you should talk to a different type of expert altogether. I recommend you talk to Dr. X who has a great deal of knowledge in that field."

A third important attribute of pyramiding we think worthy of further study is that it allows (or invites) referrals to people *outside* a predefined population. We did not take advantage of this attribute of pyramiding during the research we reported upon here. However, if we had done so, it is likely that the efficiency of the pyramiding process would have been still further improved. For example, if we had asked our group members, "Who in this group *or elsewhere* do you think has more jazz CDs than you do?" we probably would have found many additional promising target individuals. Screening does not offer researchers a similar opportunity – it is restricted to a predefined population. The ability of pyramiding to cross population boundaries has proven very valuable in practice. For example, in lead user studies it has been found that the "best" individuals are often not located within the population originally explored, but rather are often located in fields distant from the target market under study (Lilien et al. 2002, Franke and Pötz 2008).

All in all, we think that a great deal of interesting and valuable research can be done on the subject of pyramiding, and we hope that others will join us in further exploration of this method.

References

Atkinson, R. and Flint, J. (2001), "Accessing Hidden and Hard-to-Reach Populations: Snowball Research Strategies," *Social Research Update* 33. Guildford: University of Surrey. http://www.soc.surrey.ac.uk./sru/SRU33.html

Biernacki, P.; Waldorf, D. (1981), "Snowball Sampling: Problems and Techniques of Chain Referral Sampling," *Sociological Methods and Research* 10 (2), 141-163.

Burt, R.S. (1992). *Structural Holes: the Social Structure of Competition*, Cambridge, MA: Harvard University Press.

Dodds, P. S., Muhamad, R., and Watts, D. J. (2003). "An Experimental Study of Search in Global Social Networks," *Science*, 301 (6534), 827-829.

Franke, N., and Pötz, M. (2008). "The Analogous Market Effect: The Contribution of Users from Analogous Markets to the Process of Idea Generation," Working Paper Vienna University of Economics and Business Administration.

Goodman, L. (1961). "Snowball Sampling," *Annals of Mathematical Statistics*, 32 (1), 117–151.

Guiot, J.M. (1976). "A Modification of Milgram's Small World Method," *European Journal of Social Psychology,* 6 (4), 503-507.

Hidi, S. and Renninger, K. A. (2006). "The Four-Phase Model of Interest Development," *Educational Psychologist*, 41 (2), 111-127.

Kleinberg J.M. (2000). "Navigation in a small world," *Nature,* 406 (2000), 845.

Krapp, A. (2002). "An Educational-Psychological Theory of Interest and its Relation to Self-Determination Theory," In E. Deci & R. Ryan (Eds.), *The Handbook of Self-Determination Research* (405-427). Rochester: University of Rochester Press.

Lavin, D.E. (1965). *The Prediction of Academic Performance: A Theoretical Analysis and Review of Research*, New York: Russell Sage Foundation.

Lilien, G., Morrison, P., Searls, K., Sonnack, M. and von Hippel, E. (2002). "Performance Assessment of the Lead User Idea Generation Process for New Product Development," *Management Science,* 48 (8), 1042-1059.

Lüthje, C. (2000). *Kundenorientierung im Innovationsprozess: eine Untersuchung zur Kunden-Hersteller-Interaktion auf Konsumgütermärkten*, Wiesbaden: Gabler.

Milgram, S. (1967). "The Small World Problem," *Psychology Today*, 1 (1), 61-67.

Newman, M. E. J. (2003). "The Structure and Function of Complex Networks," *SIAM Review,* 45 (2), 167–256.

Olson, E.L. and Bakke, G. (2001). “Implementing the Lead User Method in a High Technology Firm: A Longitudinal Study of Intentions versus Actions,” *Journal of Product Innovation Management,* 18 (6), 388−395.

Reagans, C. and McEvily, B. (2003). “Network Structure and Knowledge Transfer: The Effects of Cohesion and Range,” *Administrative Science Quarterly*, 48 (2), 240-267.

Renninger, K.A., Ewen, E., and Lasher, A.K. (2002). “Individual Interest as Context in Expository Text and Mathematical Word Problems,” *Learning and Instruction*, 12 (4), 467-490.

Spreen, M. (1992). “Rare Populations, Hidden Populations, and Link-Tracing Designs; What and Why?” *Bulletin Methodologie Sociologique* 36, 34-58.

Strogatz, S. H. (2001). “Exploring Complex Networks,” *Nature*, 410, 268–276.

Sudman, S. (1985). “Efficient Screening Methods for the Sampling of Geographically Clustered Special Populations,” *Journal of Marketing Research* 22 (1), 20-29.

Thomke, S., von Hippel, E. and Franke, R. (1998). “Modes of Experimentation: An Innovation Process and Competitive Variable,“ *Research Policy*, 27 (3), 315-332.

Travers, J. and Milgram, S. (1969). “An Experimental Study of the Small World Problem,” *Sociometry*, 32 (4), 425-443.

Von Hippel, E., Thomke, S. and Sonnak, M. (1999). “Creating Breakthroughs at 3M,” *Harvard Business Review***,** 77 (5), 47-57.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*, Cambridge: Cambridge University Press.

Welch, S. (1975). “Sampling by Referral in a Dispersed Population,” *Public Opinion Quarterly*, 39 (2), 237-245.