Research Policy 38 (2009) 1397–1406

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

"Pyramiding: Efficient search for rare subjects"

Eric von Hippel[a,*], Nikolaus Franke[b], Reinhard Prügl[c,d]

[a] MIT Sloan School of Management, E52-556, Cambridge, MA 02141, USA
[b] WU Wien (Vienna University of Economics and Business), Institute for Entrepreneurship and Innovation[1], Vienna User Innovation Research Initiative[2], Austria
[c] Zeppelin University Friedrichshafen, Chair of Innovation, Technology and Entrepreneurship, Germany
[d] Innsbruck University School of Management, Department of Strategic Management, Marketing, and Tourism, Austria

ARTICLE INFO

Article history:
Received 13 October 2008
Received in revised form 11 July 2009
Accepted 21 July 2009
Available online 1 September 2009

Keywords:
Sequential search
Research method
Lead user innovation
Serial experimentation
Network theory

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.

The search strategy of "pyramiding" is a potential solution to this problem under many conditions. Pyramiding is a search process based upon the idea that people with a strong interest in a topic or field tend to know people *more* expert than themselves. In this paper we report upon four experiments empirically exploring the efficiency of pyramiding searches relative to mass screening. We find that pyramiding on average identified the *most* expert individual in a group on a specific topic with only 28.4% of the group interviewed – a great efficiency gain relative to mass screening. Further, pyramiding identified one of the top 3 experts in a population after interviewing only 15.9% of the group on average. We discuss conditions under which the pyramiding search method is likely to be efficient relative to screening.

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 the subset with 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 would clearly be beneficial.

One method 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 search is a variant of snowballing – but with an important difference. Pyramiding requires 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 search assumes 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 high 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 search 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

* Corresponding author.
E-mail addresses: evhippel@mit.edu (E. von Hippel), nikolaus.franke@wu.ac.at (N. Franke), reinhard.pruegl@zeppelin-university.de (R. Prügl).
[1] http://www.e-and-i.org.
[2] http://www.userinnovation.at.

0048-7333/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.respol.2009.07.005

the process until individuals with the desired high levels of the attribute have been identified.

Pyramiding has already proven its usefulness in studies seeking lead users within a population of product users. Lead users 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 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 mass screening method. However, lead users are relatively rare in a population, and so screening can be quite inefficient. For example, Lüthje (2000) reported screening 2043 persons to identify 22 lead users in a lead user study – a sampling efficiency of only 1.1%. Eager to avoid low efficiencies such as these, researchers conducting lead user studies have increasingly turned to the pyramiding search method to achieve more efficient identification of lead users (e.g. von Hippel et al., 1999; Olson and Bakke, 2001; Lilien et al., 2002).

Even though researchers seeking efficiency in finding rare subjects increasingly turn to pyramiding, the actual efficiency of pyramiding efficiency vs. mass screening has never been empirically tested. Clearly, it is important to do this if pyramiding is to become a well-understood 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 mass screening search methods, and report upon an informal pilot study comparing the two methods. In Section 3 we review related literature. In Section 4 we report upon our study of 663 pyramiding search chains in 18 search settings and compare the efficiency of these with mass screening methods applied to the same settings. In Section 5 we discuss our findings, and discuss the real-world conditions under which pyramiding is likely to be a more efficient search method than mass screening.

2. Pyramiding vs. screening searches

2.1. Background on pyramiding and screening

Pyramiding search, as was mentioned earlier, is a variant of "snowball sampling" (Goodman, 1961; Welch, 1975). Snowballing assumes that people in any population will tend to personally 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 to identity 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).

Pyramiding (and snowballing) differs from mass screening in that it applies a questionnaire to a group of people in series – in essence it is an experimental protocol involving *x* experiments conducted in series. Mass screening, in contrast, is an experimental protocol involving *n* experiments conducted in parallel. 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 (e.g. Loch et al., 2001). In the case of experiments conducted in parallel, no learning is possible between experiments.

Pyramiding applies its series of experiments to conduct "hill-climbing" – a 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 encountered (Rivkin and Siggelkow, 2002; Siggelkow and Rivkin, 2005, 2006). A well-known disadvantage of the hill-climbing strategy is that the researcher will not know whether the highest point reached is in fact the highest peak on the landscape or simply a foothill.

Although the search strategy of "pyramiding" involves the hill-climbing metaphor as just described, in the case of pyramiding each location on the hill reached by the researcher is not just a physical point in the landscape, but an intelligent actor (e.g. a person or an

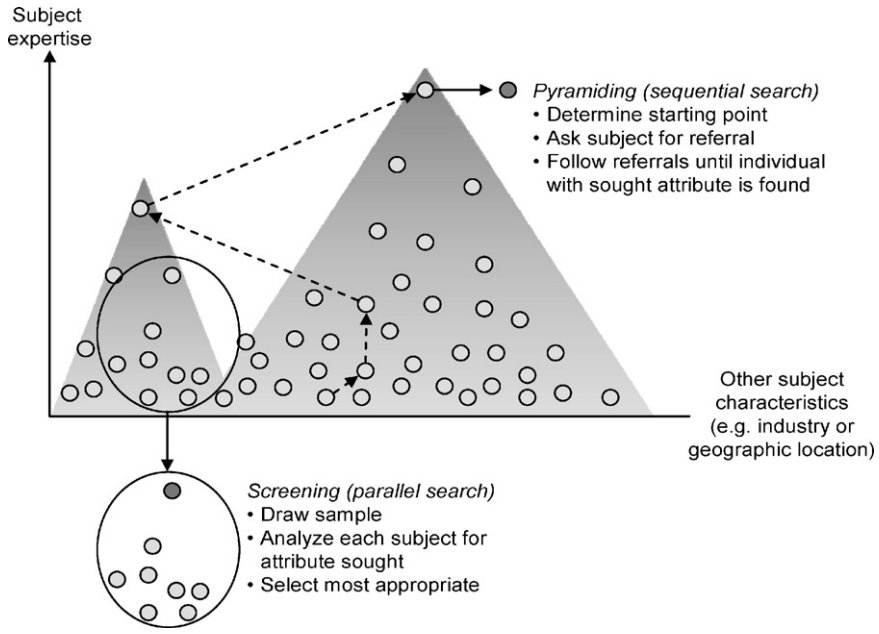

Fig. 1. The search concepts of screening and pyramiding.