J PROD INNOV MANAG 2015;••(••):••–••

DOI: 10.1111/jpim.12291

The Social Network Position of Lead Users*

Jan Kratzer, Christopher Lettl, Nikolaus Franke, and Peter A. Gloor

The field of lead user research has seen a great deal of attention from academics and practitioners alike. However, we still lack a full understanding of the nature of users with high potential for innovation. In this paper, we employ a social network perspective on lead users. Increasing the realism of our research in three empirical studies with different empirical settings and methods, we provide robust evidence that lead users have a distinctive social network position: They exhibit an unusually high level of "betweenness centrality," meaning that they are positioned as bridges between different social groups. This finding has two major implications for lead user theory. First, it consolidates seminal conceptual work on lead users and their embeddedness in social networks. And second, the findings extend and validate prior work on the social network perspective of lead users by combining theoretical insights from cognitive psychology, research on creativity, and network theory. As the social network positions of individuals can be mapped quickly and at low cost with modern Web mining tools, our findings may point to a new and readily applicable approach for the efficient and effective identification of lead users in real-life projects, an aspect that is usually emphasized as the most crucial activity in lead user projects.

Practitioner Points

• Lead users can be identified as bridges in (online) social networks using readily applicable software tools.
• A social network perspective of lead users can be a fruitful complement to other lead user identification methods.
• As lead users represent bridges in (online) social networks, they are not only relevant as lead users but for all attempts in distributing information through social networks.

Introduction

The lead user approach has been shown to be one of the most effective methods to identify ideas and concepts for really new products with high commercial attractiveness (e.g., Franke, von Hippel, and Schreier, 2006; Herstatt and von Hippel, 1992; Kratzer and Lettl, 2009; Lilien, Morrison, Searls, Sonnack, and von Hippel, 2002; Lüthje, Herstatt, and von Hippel, 2005; Morrison, Roberts, and von Hippel, 2000; Urban and von Hippel, 1988). von Hippel (1986, p. 796) defines lead users based on two characteristics: "Lead users face needs that will be general in a marketplace—but face them months or years before the bulk of that marketplace encounters them, and Lead users are positioned to benefit significantly by obtaining a solution to those needs." The extent to which an individual is a lead user is domain-specific (von Hippel, 1986, 2005). Consequently, the construct is referred to as the lead userness of an individual in a specific domain, and its measurement comprises two dimensions: trend leadership and high expected benefit. Individuals considered to be lead users score high on both dimensions (Franke et al., 2006; von Hippel, 1986, 2005).

However, integrating lead users into new product development requires the ability to identify them effectively and systematically. Prior research has emphasized that this is the most critical issue in this approach (Bilgram, Brem, and Voigt, 2008; Lüthje and Herstatt, 2004). In early studies on the topic, lead users were often sought using a "mass screening" approach. This means that the firm collects information on every member of a population (e.g., from the customer database) and determines who displays the most lead user characteristics. However, as lead users are quite rare, screening is not very cost-effective. Therefore, in recent years, the search method of "pyramiding" has gained popularity (von Hippel, Franke, and Prügl, 2009). In this approach, the researcher asks individuals to identify others who they think have attained a higher level of lead user characteristics than they have themselves. The researcher then

Address correspondence to: Jan Kratzer, Entrepreneurship and Innovation Management, Center for Entrepreneurship, Technical University Berlin, Strasse des 17. Juni 135, Berlin 10623, Germany. E-mail: jan.kratzer@tu-berlin.de. Tel: 0049-3-314-26581.

*The authors acknowledge the help of Florian Schirg, Susanne Roiser, Ilse Klanner, Christopher Hader, Stephanie Ambach, Isabella Gaupmann, and Christoph Sohn.

poses the same question to the persons identified and continues the process until individuals with the desired high levels of lead user characteristics have been found. Still, the effort involved in this process is considerable and can make up a large percentage of total costs (Moriarty and Bateson, 1982).

The advent of online social network applications might constitute a new and potentially far more effective avenue for lead user identification. In virtual social networks, users exchange information and collaborate online, thus leaving a mark on those networks. Based on a number of studies, this study has the underlying assumption that there are also lead users among the masses of users active on the Web (Füller, Matzler, and Hoppe, 2008; Jeppesen and Laursen, 2009; Long and Siau, 2007; Sawhney, Verona, and Prandelli, 2005). This assumption is also supported by the great interest shown by Lego (Berg-Jensen, Hienerth, and Lettl, 2010), Dell (Cooke and Buckley, 2008), Procter & Gamble (Huston and Sakkab, 2006), and many other leading companies with respect to the innovation potential of online communities.

In this paper we are particularly interested in whether the social network perspective on lead users as introduced by Kratzer and Lettl (2008, 2009, 2011) can be leveraged for lead user identification. If there are patterns in the social network position of lead users, then their position might serve as a shortcut in identifying them. Information on the connectedness of persons in online social networks and communities is available online at virtually no cost, and software that computes the social position of users is becoming increasingly powerful and easy to use. Compared with today, therefore, lead user identification might become a far quicker and less expensive process. This study provides the next step in this line of research and investigates the fundamental research question underlying these new ways of lead user identification: *Do lead users have a distinct social network position that allows their identification?*

On the basis of prior research on lead users, insights on problem solving from cognitive psychology, as well as research on creativity and network theory, hypotheses on the distinctive social network position of lead users are developed. In order to test these hypotheses, three studies in different settings are conducted. This multistudy approach was applied in order to increase the realism of this research and to verify the robustness of the findings. This objective could not have been attained by conducting a single, integrated study.

In Study 1, a unique data set of $n = 267$ adolescents in 11 high school groups is used and finds that lead userness is indeed systematically linked with a specific type of network position. In Study 2, the realism is increased by comparing the network position of 126 real-world lead users (identified through costly pyramiding approaches) with that of 141 nonlead users on the basis of data from 15 recent lead user studies conducted by 11 companies in different industries (Deutsche Telekom, Frequentis, I.S.A., Kotányi, MAM Babyartikel, OMV, Ottakringer, Palfinger, Schindler, Siemens, and Stock Austria). The findings are again robust. Finally, Study 3 tests the instrumentality of the findings by analyzing the extent to which individuals identified on the basis of their specific network position are able to provide valuable input to an ongoing lead user study conducted with another company (Airbus).

Overall, the findings clearly indicate that lead users do have a distinct social network position: They are positioned as bridges between different local groups. Therefore, the "betweenness centrality" of an individual is an effective indicator of his or her lead user status. As this

BIOGRAPHICAL SKETCHES

Prof. Dr. Jan Kratzer is the chair of entrepreneurship and innovation management and the managing director of the Center for Entrepreneurship at Berlin Institute of Technology, Germany. He holds a doctoral degree from the Interuniversity Center for Social Science Methodology and theory (ICS) in the Netherlands. His research is mainly on factors that drive entrepreneurial activities toward success, and social networks as phenomenon determining all activities and hierarchies of business life.

Dr. Christopher Lettl is a full professor and the director of the Institute for Strategy, Technology and Organization at the Vienna University of Economics and Business. He holds a doctoral degree from the Hamburg University of Technology and a Venia Legendi (Habilitation) from the Berlin University of Technology. His research takes point of departure in the following paradigm-shifting phenomenon: the shift from closed and producer-centered toward open and user-centered innovation processes. In his research, he analyzes how organizations may best manage the opportunities and threats induced by this phenomenon with a specific focus on implications for strategy and organization design.

Prof. Dr. Nikolaus Franke is a full professor and the director of the Institute for Entrepreneurship and Innovation at the Vienna University of Economics and Business. In his research, he analyzes methods and instruments that allow companies to benefit from open and user innovation, such as lead user analysis, toolkits for user innovation and design, or crowdsourcing.

Dr. Peter A. Gloor is a research scientist at the Center for Collective Intelligence at MIT's Sloan School of Management, where he leads a project exploring collaborative innovation networks. He is also the founder and chief creative officer of software company galaxyadvisors, an honorary professor at the University of Cologne, a lecturer at Aalto University in Helsinki, and a distinguished visiting professor at P. Universidad Católica de Chile. Earlier he was a partner with Deloitte and PwC, and a manager at UBS. He got his Ph.D. in computer science from the University of Zurich and was a post-doc at the MIT Lab for Computer Science.

information can be retrieved easily from the Web, it might be possible to make the identification of lead users easier, faster, and more cost-effective.

Development of Hypotheses

A number of recent studies have explored the network position of lead users. For example, Franke and Shah (2003) examined the characteristics of lead users in a study of four offline extreme sport communities. They found that lead user innovators were significantly more active in their communities than average users and systematically used input from others to develop their innovations. Belz and Baumbach (2010) confirmed the central position of lead users by proceeding in the opposite way: They identified the 40 most active members in an online forum and then analyzed the extent to which they exhibit lead user characteristics. Their findings likewise indicate that there is a systematic relationship between a user's leading-edge status and his or her social network position. But what exactly is this relationship? Lead users and average users might differ in two important dimensions: (1) their degree centrality, that is, the number of direct contacts an individual has within a social network, and (2) their betweenness centrality, that is, their position as a bridge between different social groups.

Degree Centrality of Lead Users

Degree centrality is characterized as the number of contacts directly connected to the focal actor.

Lead users are defined as having needs that cannot be satisfied by existing products on the market. The main incentive for lead users to innovate is therefore the need to find appropriate solutions for their own needs (von Hippel, 1986). Developing an initial idea into a functioning prototype will usually exceed the capabilities of the individual, making it likely that such users will seek the assistance of others. Peers can provide valuable feedback on ideas or solutions, and this "peer review" process enables innovative users to advance and improve their ideas (Franke and Shah, 2003). Peers facilitate a process of collaborative filtering and rigorous selection for promising ideas (Brown and Duguid, 2001; Perry-Smith and Shalley, 2003; von Hippel, 2007), thus making it possible to filter out ideas that have no potential to become an effective solution. Innovative users who are centrally embedded among their peers are thus able to make better choices as to which ideas and novel combinations are worth advancing (Fleming, 2007).

Peers might provide helpful input in the development phase, as an innovative solution often requires access to tacit knowledge (Polyani, 1966). This kind of knowledge is most likely to be exchanged via direct contacts and not indirectly (Nonaka, 1994; Weimann, 1982). Direct contacts enable the mutual development of local orientation and coding schemes (Wilensky, 1967), which then make it possible to translate contextual cues into pure product-related information. This argument feeds the expectation to find a relationship between an individual's lead userness and his/her number of direct contacts within a social network.

However, this relationship is likely to be nonlinear. Beyond a certain boundary, the marginal value of additional direct contacts with peers may become negative. This is the case because the benefits arising from more contacts will converge to a boundary value, while their costs and negative effects will increase in a strictly monotonic way. The likelihood of redundancy increases with the number of direct contacts. The difference between one and ten contacts may be dramatic, but the difference between 1000 and 1010 contacts is likely to be minimal; in fact, it will most likely be zero. Establishing and maintaining contacts requires time and effort. Increasing the number of contacts will increase the likelihood of attentional overload, which limits cognitive focus and may cause individuals to rely on cognitive shortcuts such as heuristics and preexisting schemata (Shalley, 1995). Theories on the phenomenon of "creativity blocking" suggest that high levels of interaction may carry individuals along by the momentum of their enthusiasm for an innovative idea rather than a clear evaluation of its real value. Such unbridled enthusiasm can lead to poor decisions (Nicholas, 1994; West, 1990) and reduce the number and quality of solutions generated (Nyström, 1979). The net effect of benefits and costs will thus have a maximum. This suggests that there is an optimum number of direct contacts for lead users seeking information (which is, of course, rather difficult to identify).

So far, the causal direction assumed is that the lead user characteristic of having a high need for innovation will prompt the individual to seek contacts, which will then help him/her develop innovative solutions ("lead userness $\rightarrow$ centrality"). However, there are also arguments for the opposite causal direction, namely "centrality $\rightarrow$ lead userness." Given that (up to a certain maximum value) contacts can provide the individual with access to topical and valuable information (Perry-Smith and Shalley, 2003), the number of contacts will impact the trend position of the individual user: Those who know "what's going on" will exhibit a leading trend position (i.e., the second component of the lead user construct; von Hippel, 2005), which in turn augments the likelihood

that the individual will come up with a novel idea. Franke et al. (2006) termed this phenomenon the "supply side effect" of lead user theory and provide empirical evidence that this effect actually exists. Similarly, Schreier, Oberhauser, and Prügl (2007) find that access to product and use knowledge is an antecedent of lead user status. The conclusion is that there is not only one direction of causality between degree centrality and lead userness. Despite their conceptual distinctiveness, they exhibit a mutual causal relationship: High lead userness leads to high degree centrality, and high degree centrality leads to high lead userness. For the purpose of this study, which is to determine whether social network position is associated with lead userness, the direction of causality is not of primary interest. After all, effective predictions require a stable correlation between indicator and event, not necessarily a causal relationship (Spirtes, Glymour, and Scheines, 2000).

H1: There is an inverted U-shaped relationship between the degree centrality of a user in social networks and the extent to which that user displays lead user characteristics.

Betweenness Centrality of Lead Users

Betweenness centrality is characterized as the number of shortest paths that pass through a given actor in a network. This kind of centrality is not based on the number of ties, but on the extent to which an actor facilitates the flow of information by being positioned on many information paths (e.g., Borgatti, 1995).

The notion that lead users exhibit a higher degree of betweenness centrality in social networks than average users is supported by arguments from lead user research, cognitive psychology, and social network theory.

A number of studies have shown that lead users apply knowledge from different domains and disciplines in order to develop solutions to their problems. For example, the trendy sport of kite surfing is a user innovation that combines surfing and hang-gliding (Tietz, Morrison, Lüthje, and Herstatt, 2005), while the user innovation of minimally invasive surgery is a novel recombination of new surgical techniques with advanced software applications and robotics (Lettl, Herstatt, and Gemünden, 2006). Insights into the process of developing archetypal lead user innovations such as the first device for gas chromatography (Riggs and von Hippel, 1994), the first surgical navigation systems, the first medical robot for neurosurgery, the first biocompatible implant for hernia surgery (Lettl et al., 2006), and the first mountain bike (Lüthje et al., 2005) reveal that the lead user inventors combined diverse fields of technological knowledge. These examples correspond to the general pattern of innovative processes requiring a novel recombination of diverse knowledge bases (Nelson and Winter, 1982; Schumpeter, 1939; Usher, 1954).

Similarly, researchers in cognitive psychology propose that creative problem solving is stimulated by reassembling elements from existing knowledge bases in a novel fashion (e.g., Dahl and Moreau, 2002; Ebadi and Utterback, 1984; Ward, 1994). The cognitive processes propelling creativity involve the ability to recognize links between seemingly different concepts, which leads to solutions that depart from established ways of doing things (Guilford, 1950; Mumford and Gustafson, 1988) and enables individuals to come up with different types of responses to a single problem (Torrance, 1974).

Consequently, lead users will strive to gain access to as many relevant knowledge domains as possible. However, this does not necessarily involve a large number of direct contacts (degree centrality). Burt (1992) argues from a social network perspective that information access and control advantages are created when relations *bridge* groups. Actors with a social network position that connects different groups are able to monitor information more effectively, and they receive information more quickly (Staber, 2004). Being positioned as a bridge between distant social groups thus puts an individual in a position to access very different knowledge domains. Lead users will therefore aim for such a position.

Again, no one-way causality in this context is assumed. If an individual exhibits a high level of betweenness centrality, it will also impact his or her lead user position. Bridging different knowledge domains means having access to less redundant and thus more valuable information (Staber, 2004). Hardly any research has been conducted on the relationship between lead users and betweenness centrality. The only studies the authors are aware of are those conducted by Kratzer and Lettl (2008, 2009, 2011) and Kunst and Kratzer (2007), which examine a special population (young schoolchildren). Although children differ from adults in many respects and it is particularly questionable whether they are able to provide valuable input in real lead user studies, the correlations between lead userness and betweenness centrality underscore the plausibility of this reasoning.

H2: The betweenness centrality of a user in social networks is positively related to the extent to which that user displays lead user characteristics.

To test the hypotheses above, three studies in different empirical settings are conducted. The rationale underly-

ing these studies is as follows: Study 1 uses traditional social network analysis within the population of adolescents in high school groups. It serves as the basis for the two ensuing studies. Traditional social network analysis requires that the boundaries of the social entity under study can be defined. This first study is also based on self-assessments of individuals. The latter two studies aim to align the research with the real-world context of lead user studies conducted by companies in order to increase the realism of this research. In Study 2, data from 14 real-world lead user projects conducted by 11 companies in different industries is used. This study compares the individual social network positions of previously identified lead users with those of nonlead users. The contribution of this approach is that the classification of an individual as a lead user is unusually valid because it is based on an external assessment by experts who are motivated to maximize the commercial success of the respective projects. Study 3 takes a complementary approach and asks whether people identified via their social network position also display strong lead user characteristics. This perspective again adds realism to the research, as managers would ask for the instrumentality of social network information for lead user identification. The empirical setting for this study is a lead user project conducted by Airbus in close cooperation with the university where two of the authors are based.

Study 1: Evidence from Network Data

Study Design, Procedure, and Participants

In order to measure the distinctive social network position of lead users, data from 11 high school classes in the Netherlands were gathered. The participants were 267 young adults between 14 and 17 years of age, and the total response rate was 94.4%. This data set has important advantages for the purposes of the research because it allows us to measure the individuals' social network position with high validity. First, investigating high school classes allows measurement of social networks with defined boundaries. As participants spend a large part of their day at school (around six to eight hours per day in the Netherlands), full networks of high school classes (Defares, Kema, Van Praag, and Van der Werft, 1971) were investigated. Second, participants may have less individual strategic interest in showing that they have many contacts at their disposal (as many grown-ups would do), so it is assumed that degree centrality and betweenness centrality measures are largely unbiased. The average size of a group was around 24 persons, the participants were 14–17 years old ($M = 15.43$, $SD = .46$), and the gender breakdown was 47% women and 53% men.

Data Collection and Measures

All 267 participants were provided with a full roster of their schoolmates and were asked to indicate with whom they discussed problems using a given scale ("never," "not often," "sometimes," "often," and "very often"). This procedure is typical in the collection of data on communication networks (e.g., Kratzer and Lettl, 2009; Leenders, Van Engelen, and Kratzer, 2007; Sosa, Eppinger, and Rowles, 2004).

Measuring social networks. The data gathered using the full rosters resulted in 11 matrices containing directed relations between all classmates in all 11 classes. The matrices were dichotomized, taking "sometimes" as the threshold. This was necessary in order to calculate betweenness centrality as proposed by actor *i* about actor *j*, and outgoing ties, that is, indications made by individual *j* about *i*, are particularly important because they cannot be interpreted in the same way. In this study, however, there is no meaningful distinction between incoming ties and outgoing ties, as two connected individuals communicate regardless of which one sends or receives the information (e.g., Barnowe, 1975). QAP correlations[1] were conducted (Krackhardt, 1987) to measure the correlation of in-link and out-link indications. In all classes, the correlations were higher than .70 and significant ($p < .001$), and the reciprocity of ties in all classes exceeds 72%. The analyses were also conducted separately for incoming ties and outgoing ties, and no differences in the results were found. It was decided to symmetrize the matrices, which also served to increase the robustness of the data. The analyses presented are based on the symmetrized and dichotomized matrices.

Degree centrality. Formally, the degree centrality of individual *i* is described by

$$D_i = \sum_{j=1}^{n} a_{ij}, \quad i \neq j$$

[1] The quadratic assignment procedure (QAP) correlation can deal with systematic interdependencies such as those usually found in social networks. The QAP is a technique for testing statistical significance using social network data. One assumption of this statistical technique is that it determines statistical significance by comparing observed values to appropriate theoretical distributions, so that the observations being analyzed are independent of one another (Krackhardt, 1987).

where a_{ij} is the number of contacts between i and j. Degree centrality was calculated with UCINET VI (Borgatti, Everett, and Freeman, 2002).

Betweenness centrality. Formally, the betweenness centrality of individual k is described as

$$B_k = \sum_i^n \sum_j^n \frac{g_i k_j}{g_{ij}}, \quad i \neq j \neq k.$$

Specifically, g_{ij} is the number of geodesic paths from i to j, and g_{ikj} is the number of these geodesics that pass through individual k (Freeman, 1979; Wasserman and Faust, 1994). To put it in less technical language, information traveling from one individual to another can take different routes through the network, and some routes are longer (i.e., pass through more individuals before reaching the intended target) than others. For every possible combination of a sender and a target, there is at least one route that is shortest, and for the sake of simplicity it is assumed that information tends to take the shortest route from the sender to the intended target. Individuals who tend to be part of many such shortest routes then have a considerable information advantage over those who are rarely located on these information routes. The sum of the proportions of all shortest routes between any two people in the network that pass through person i is called betweenness centrality. UCINET VI was used to calculate this measure (Borgatti et al., 2002).

Measuring lead userness. Lead userness is a domain-specific construct (von Hippel, 1986, 2005). It is measured by six indicators that are based on existing scales (Franke et al., 2006; Morrison et al., 2000; Morrison, Roberts, and Midgley, 2004). A Likert-type scale of 1 to 5 (1 = never; 5 = very often) is used. The two lead user characteristics as defined by von Hippel (1986) are the basis for measurement. The first characteristic "trend leadership" is measured by items 1, 2, and 3. The second characteristic "high expected benefit" is measured by items 4, 5, and 6.

The domain for this study is Hyves, an online social network application that was very popular in the Netherlands at the time of this study. According to the calculations of STIR (Stichting Internet Reclame, 2009), the platform had attained a reach level of 80.5% in the 13–20 age group at that time. Hyves provides the opportunity to develop open social applications and introduce them as "gadgets" to other Hyves users.

The questions were adapted to the setting as follows.

1. I often find out about new Hyves applications earlier than others.
2. I am always up to date concerning Hyves applications.
3. I have already enjoyed added value by being one of the first to use new Hyves applications.
4. With regard to Hyves applications, I have requirements that are not covered by the market.
5. I am often disappointed with bad Hyves applications.
6. In my view, most Hyves applications need to be improved.

The scale reached a Cronbach's alpha of .82. Thus, the indicators can be combined into a single measure of lead userness due to sufficient internal consistency. Confirmatory factor analysis was applied in order to assess the quality of the measurement. There is good overall fit (goodness of fit index [GFI] = .98; incremental fit index [IFI] = .98; comparative fit index [CFI] = .96). All indicators loaded positively and were statistically significantly on lead userness ($p < .01$). This suggests a good degree of convergent validity.

Analytical Techniques

In order to test H1 and H2 based on more reliable measures of social network positions, a latent class analysis resulted in a six-cluster solution. Accordingly, the participants were broken down into two groups, with Group A comprising the participants who scored highest in lead userness ($n = 34$) and Group B containing all other participants from the remaining five clusters combined ($n = 233$). A technical outline can be found in Appendix A. The high correlation between degree centrality and its squared term would introduce problems of multicollinearity in the analysis. In order to resolve this problem, the squared term was centered by subtracting the overall mean from it. Multicollinearity was tested by checking the variable inflation factor and condition index, as well as the distribution of residuals. These tests showed that it is feasible to include all of the variables in the binary logistic regression analyses (Neter, Kutner, Nachtsheim, and Wasserman, 1996).

Findings and Discussion

The analyses show that lead users do exhibit significantly higher betweenness centrality than nonlead users, which supports H2 (Table 1). There is no statistically significant indication of group differences concerning degree centrality. A Levene's test shows that the variances within the groups do not differ, suggesting that H1 should be rejected.

Table 1. Degree and Betweenness Centrality of Lead Users versus Nonlead Users—Study 1[a]

	Betweenness Centrality			Degree Centrality		
	n	Mean (*SD*)	*p*	*n*	Mean (*SD*)	*p*
Lead users	34	2.49 (1.03)	.000	34	2.16 (.95)	.187
Nonlead users	233	2.11 (.98)		209	2.04 (.92)	

[a] Two-tailed tests are reported; *t*-test.

The analyses show degree centrality to be unrelated to the lead user construct and must therefore discard H1, as illustrated in Table 2. Apparently, a person's number of contacts is not associated with lead user status. On the other hand, clear support for H2 was found: Betweenness centrality is associated with lead userness.

One limitation of Study 1 is the measurement of lead user status, which relied on self-assessments. Some arguments indicate that such a measurement can claim a certain degree of validity. For example, Heneman (1974) has shown that when anonymity is promised, as in this study, self-reported measures exhibited less restriction in range and leniency than more objective ratings. Others have also successfully applied self-assessments in the past (e.g., Cooper, 1981; Leenders et al., 2007). However, as it is one of the two core constructs, we attempted to provide more valid evidence by carrying out another study.

Study 2: Increasing Realism with Data from Lead User Studies

Study Design, Procedure, and Participants

Building on Study 1, Study 2 again tested the relationship between a person's centrality and his/her lead user status, at the same time attempting to overcome the limitations associated with the way of measuring the lead user construct in the previous study. In order to increase the realism of this research, data from 14 *real-world* lead user projects conducted by 11 companies in different industries were used. The individual social network positions of the lead users (i.e., individuals who had been identified as lead users in the respective projects) with those of nonlead users was compared. For the latter, the reference group consisted of individuals who were considered not to be lead users in the respective projects, yet had relevant knowledge in the respective search field. The contribution of this approach is that the classification of an individual as a lead user is based on an external assessment that involved a great deal of effort. If the hypotheses are correct, then these real-world lead users should exhibit a clearly different social network position compared with nonlead users. Regarding the measurement of network position, it was focused on betweenness centrality alone. So for two reasons these were done: First, the findings from Study 1 clearly indicate that degree centrality is not associated with lead userness. Second, measuring degree centrality using Web mining approaches is problematic when the network boundaries are unknown because the nearest nodes in a small-world Web structure might contribute very little to the importance of a node. What makes the difference is how important those neighboring nodes are, which is reflected in betweenness centrality.

Table 2. Binary Logistic Regression—Study 1[a,b,c]

	B	Wald
Constant	3.34	24.55**
Age	.05	3.27
Gender	−.01	.05
Degree centrality	.04	3.21
Degree centrality^2	−.03	1.89
Betweenness centrality	2.98	47.81**
Nagelkerke's R^2	.14	
Chi-square	94.34	

[a] Two-tailed tests are reported; *$p < .05$, **$p < .01$. [b] $n = 267$. [c] 1 = lead users; 0 = nonlead users.

Data Collection and Measures

Measuring lead userness. All the lead user projects were conducted in close cooperation with the university where two of the authors are based. The explicit goal of the projects was to develop radically new product concepts and solutions that are attractive to customers. They were structured in four phases: (1) definition of problem, objectives, and constraints; (2) identification of relevant needs and trends; (3) identification of lead users based on the most important trends; and (4) idea generation and concept design at a three-day lead user workshop (see Lilien et al., 2002; Lüthje and Herstatt, 2004; Urban and von Hippel, 1988). The projects required roughly four months of work on average, and the companies were highly satisfied with the outcome in all cases.

For this research purposes, the third phase of the projects was relevant. Here, the lead users were identified in the target market as well as in other comparable markets. In all projects, this was done through a combination of pyramiding, content analysis, and problem broadcasting in online forums.

The final decision to qualify a person as a lead user (or not) was based on intensive interviews in which the candidates' (1) trend positions, (2) need for individual innovations and innovation track record, and (3) ideas and possible solutions to the specific problem in question were thoroughly examined by trained project members. The effort required to identify the 140 lead users was considerable, comprising 3118 interviews and searches in over 1300 online forums. These huge investments have two important implications: First, they support the initial proposal that it would be helpful to gain a better understanding of how to reduce the costs of searching for lead users. Second, they suggest that the identification of lead users was quite valid, far more so than the self-assessments used in the previous study.

The social network position for individual lead users (i.e., individuals who had been identified as lead users in the respective projects) and individuals from the respective reference groups (i.e., individuals who were not considered to be lead users in the respective projects, yet had relevant knowledge in the respective search field) were computed and compared using the Condor software program. A decision was made to use two different reference groups in order to determine whether lead users actually have a different social network position compared with nonlead users.

In one project (Deutsche Telekom), the reference group of users who had qualified as candidates during the search phase but were assessed as having less pronounced lead user characteristics than the small group of candidates who were ultimately classified as lead users was used. This means that users in this reference group were assessed with lower scores on trend leadership, expected benefit, and development of ideas/solutions for the specific problem compared with users in the lead user group. Note that this reference group of nonlead users is not neutral and must still be considered to have far stronger lead user characteristics than average users. After all, they had been selected as promising candidates in the lead user search process.

In the other projects, a different reference group is used, namely those 57 company employees (two to four per project) who conducted the lead user study on the company side and worked together with the lead users on new solutions during the lead user workshop (Phase 4). These employees were highly qualified managers with various functions within their companies (usually research and development, marketing, and production). The reasoning is that this group of nonlead users will have a similar interest in the given problem as the 140 lead users, but from a different perspective: The former obtain benefits from *selling*, not from *using* a new solution. According to the literature on user innovation, this decreases the likelihood of coming up with radical new ideas (von Hippel, 1986, 2005) and makes them a meaningful reference group.

Again, taking them as a point of comparison is quite demanding because they are particularly well qualified in the search field of the lead user study (i.e., product knowledge, process knowledge, knowledge about the value chain and competition) and have considerable experience in innovation management. This makes the study rather conservative. If significant differences are observed between these groups, the effects with regard to average users should be even stronger. Table 3 provides an overview of the studies and the firms involved.

Measuring betweenness centrality. The betweenness centrality of individuals in the discussion network of online expert discussion forums or Web sites is measured. A software tool to look at who discusses which topics, as this reflects the innovativeness of the person's contribution, is used. A network by drawing a connection between people if they were linked back in their contributions to the forums and other Web sites is constructed. In order to measure the betweenness centrality of the subjects in the sample (lead users as well as nonlead users), Condor is used, which enabled us to compute this variable for a given person quite easily. The advantage of such methods is that they are also easy to use in practical applications (i.e., real-life lead user searches). Condor is based on a simple idea: "You are who links to you." Condor applies a process called a "degree-of-separation search," which constructs a graph where the nodes are either the search terms, or the Web sites where the search terms occur, using high betweenness forums and Web sites returned in a search engine query for a name as a proxy for the significance of this person (Frick, Guertler, and Gloor, 2013; Gloor, Krauss, Nann, Fischbach, and Schoder, 2009). For example, Condor collects the most important Web pages mentioning "Joe Smith" and then inputs these URLs into a Web search engine to see which other Web sites link back to them. This process leads to a network map (Figure 1) that displays the search term "Joe Smith" at its core, surrounded by the Web sites or blog posts returned in response to the search query, or the links among posters responding to an original post in an online forum (Gloor et al., 2009). In Figure 1, the Web pages containing the search results are at the end of the black lines originating from the search term "Joe Smith." The dark gray lines at the end of the black lines connect the URLs linking back

Table 3. Overview of Lead User Studies Used in Study 2

Company	Industry	Number of Interviews for Lead User Identification	Number of Forums Searched	Lead Users Identified
Deutsche Telekom	Telecommunications	205	498	18
Deutsche Telekom	Telecommunications	240	n.a.	9
Frequentis	Airline communications and information	367	n.a.	10
I.S.A.	Automation	270	28	7
Kotányi	Food	220	162	9
MAM Babyartikel	Consumer goods	246	137	9
OMV	Energy	60	n.a.	7
Ottakringer	Food	250	40	10
Palfinger	Forklift trucks	160	200	9
Palfinger	Forklift trucks	157	184	9
Schindler	Escalators	223	10	10
Schindler	Escalators	206	99	8
Siemens	Automotive	158	n.a.	9
Siemens	Communications technology	170	n.a.	8
Stock Austria	Food	186	n.a.	8

to the Web pages containing the search term "Joe Smith."

Degree-of-separation searches are a practical way to find the most influential nodes in a given subset on the Web. By combining the graphs—the search term surrounded by the Web linking structure—returned by different degree-of-separation searches, the betweenness centrality of different individuals and identify those with the highest betweenness values can be compared. Those individuals represent bridging links on the Web or in the blogosphere (Gloor et al., 2009). In Figure 1, we see the graphs constructed for the hypothetical lead users "Joe Smith," "Mark Schwarz," "Dave Bowie," etc. This "small-world" networking structure explains the comparatively high betweenness values of lead user names in Table 4, as the search terms—the proxy for the

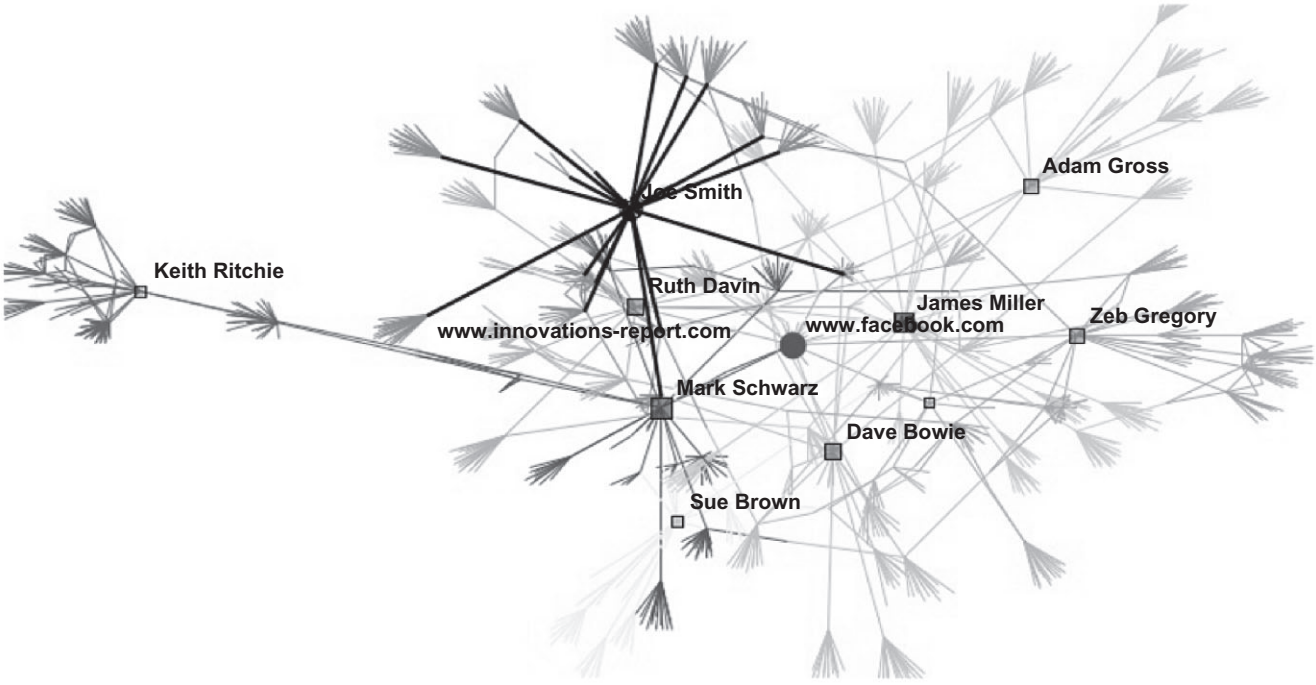

Figure 1. Sample Condor URL Network for 10 Users

Table 4. Betweenness Centrality of Lead Users versus Nonlead Users

	Sample 1[a]			Sample 2[b]		
	n	Mean (*SD*)	*p*[c]	*n*	Mean (*SD*)	*p*[d]
Lead users	16	.20 (.14)	.02	110	.21 (.20)	.000
Nonlead users	95	.12 (.13)		46	.10 (.13)	

[a] Deutsche Telekom project; nonlead users = candidates eventually not classified as lead users; standardized values. [b] Pooled sample of Deutsche Telekom (two projects), Frequentis, I.S.A., Kotányi, MAM Babyartikel, OMV, Ottakringer, Palfinger (two projects), Schindler (two projects), Siemens (two projects), Stock Austria; standardized values. [c] Two-tailed tests are reported; Mann–Whitney test. [d] Two-tailed tests are reported; *t*-test.

importance of a lead user—connect all the forums or Web sites where the name of the lead user occurs.

In this study, two limitations of Web mining tools like Condor need to be taken into account: First, it was not possible to compute values for all individuals. In some cases, there were problems with name spellings—the persons' names simply could not be found on the Web. This reduced the two samples from 136 to 111 (Sample 1) and from 197 to 156 (Sample 2) (thus explaining the figures in Table 4). It is reasonable to assume that these losses are neutral and do not influence the results. Second, only information from the Web is used. Although it is justified to assume that most interpersonal links are documented here, note that there are also offline links between individuals that are not visible on the Web.

Findings and Discussion

There is again clear support for H2 (Table 4). When measured using hard data instead of self-assessments, lead users still show significantly higher betweenness centrality than nonlead users. The effect sizes are relatively large.

Note that "nonlead user" is a relative term and must not be confused with "average" user in the context of these projects. The nonlead users in Sample 1 had been candidates for selection as lead users for good reasons—they stood out in forums, were authors of research reports and journals in the field in question, or had otherwise caught the attention of those in charge of identifying lead users. The nonlead users in Sample 2 were qualified and experienced innovation managers in the companies undertaking the lead user studies, meaning that the comparison is conservative in nature. Relative to average users, both of the nonlead user groups consist of individuals who are highly qualified for the purposes of the lead user studies. The fact the considerable differences can be seen as evidence that lead users display an unusually high level of betweenness centrality—which in turn allows us to conclude that searching for users with a high level of betweenness centrality might be an effective shortcut in lead user identification. The finding also suggests that standard software such as Condor can be used effectively to calculate the centrality of given individuals and might therefore help companies identify lead user candidates quickly and at low cost. On the other hand, this conclusion appears somewhat premature. The finding that the lead users identified exhibit high betweenness centrality does not necessarily mean that users who display high betweenness centrality are lead users. To investigate the instrumentality of the social network position as a shortcut to identify lead users, a third study was conducted.

Study 3: Exploring the Instrumental Value of Betweenness Centrality for Lead User Identification

Study Design, Procedure, and Participants

Managers seeking to identify lead users are interested in the instrumentality of the pattern found: Are users who exhibit high degrees of betweenness centrality in relevant social networks likely to be lead users capable of providing invaluable contributions to a lead user project? Building on Studies 1 and 2, a third study to test whether the findings so far pass this reality check was conducted.

The empirical setting is a lead user project conducted by Airbus in close cooperation with the university where two of the authors are based. In this project, Airbus aimed to leverage the creativity of lead users in order to develop novel solutions that increase comfort and hygiene in aircraft lavatories. In the search process for lead users, the online community PlumbingForum.com was identified as a social network in which one might find lead users with regard to hygiene and lavatories. PlumbingForum.com consists of approximately 7000 individuals who share knowledge, experience, and ideas on the topic of plumbing. Its members include expert plumbers, master plumbers, kitchen and bath designers, architects, engineers, and homeowners. PlumbingForum.com is a vibrant community that has produced more than 430,000 written messages ("posts") in over 115,000 threads. Who among the bulk of users is a lead user and might provide helpful input to the project? It would take years to study the posts in detail. Following the argument that betweenness centrality is an effective shortcut, we therefore employed Condor (see Study 2) to measure the betweenness centrality of forum members. Then it was tested whether the

50 forum members with the highest betweenness centrality values showed stronger lead user characteristics and would provide more valuable input to the project than a random sample of 150 users from the remaining forum members.

Data Collection and Measures

Measuring betweenness centrality. In order to collect data on betweenness centrality, 431,257 posts from all 13,287 individuals active in the forum were used. The link network to calculate the social network position of individuals on PlumbingForum.com by parsing the text in the forum was created. A link between two individuals is defined in two ways: First, each individual who responds to a post is linked to the previous poster. Second, the initial creator of a new thread gets a back-link from everybody who responds to the initial post. Then the full network by including all links between all individuals was compiled. The output of this social network analysis is a value of betweenness centrality for any given individual user.

Measuring lead userness. In order to determine each user's degree of lead userness, the method of netnography (Kozinets, 2002, 2010) was applied. Netnography adapts ethnographic research methods such as observation to study communication in online communities and was recently proposed as a method for lead user identification (Belz and Baumbach, 2010). In line with the previous studies, a category system (Krippendorff, 2004; Neuendorf, 2002) for the two lead user characteristics of trend leadership and high expected benefits was developed. Trend leadership using four criteria was measured: (1) the degree to which a post makes reference to state-of-the-art/new products or trends, (2) the degree to which a post provides solution-related knowledge, (3) the degree to which a post provides product-related knowledge, and (4) the degree to which a post provides advice to forum peers. High expected benefit was measured using the following two criteria: (1) the degree to which a post identifies or articulates a problem, and (2) the degree to which a post articulates dissatisfaction with current products. In addition, the extent to which a post revealed a novel solution as a validation criterion for the lead user scale is used. Two independent coders assessed each user's posts on the basis of these seven criteria. Each criterion was measured on a 3-point scale (0 = does not apply, 1 = partly applies, 2 = fully applies). One post could yield ratings for different criteria. For example, a comment containing product knowledge posted in response to a request for advice from a peer yielded ratings for the product-related knowledge criterion as well as the advice criterion. With respect to the "novel solution" criterion, only the highest rating was recorded. Therefore, a user with one very novel solution would receive a rating of 2 for the "novel solution" criterion, whereas a user with 10 partly novel ideas would only get a 1. A corresponding coding table with examples is provided in Appendix B. In order to ensure a common understanding between the two independent coders with respect to the meaning of each criterion, they were trained and provided by a number of classification examples prior to the coding process. Each coder then conducted the analysis separately.

The basis of analysis included all posts of the 200 users selected from the forum. Before analyzing the posts, the order of the sample was randomized to ensure that none of the two subsamples was biased by a primacy or recency effect (Healy, Havas, and Parkour, 2000). For all users with more than 90 posts, 90 posts for analysis were randomly selected. For users with 90 posts or less, the coders analyzed all posts. Upon completion of the content analysis, inter-coder reliability using Krippendorff's alpha was calculated and we found a satisfactory value of .72 (Hayes and Krippendorff, 2007).

Findings and Discussion

The findings clearly support the reasoning that measures of a user's betweenness centrality within a forum provide a valid shortcut for his or her identification as a lead user. Table 5 shows that those users identified as having the highest betweenness centrality in the forum clearly

Table 5. Lead Userness of Users with Top Betweenness Centrality versus Average Users

	Lead Userness[a]			Development of Novel Solutions[b]		
Population	Mean	*SD*	p^c	Mean	*SD*	p^c
Top 50 users in terms of betweenness centrality ($n = 50$)	4.98	2.86	.000	.13	.43	.036
Average users ($n = 150$)	.01	.12		.00	.00	

[a] Summed index value of seven scales. [b] Single scale. [c] Two-tailed tests are reported; *t*-test.

exhibit higher lead userness ($M = 4.98$, SD 2.86) than the random sample of average users, who have a score of $M = .01$ ($SD = .12$). The difference is significant ($p < .000$).

Does this mean that the users with high betweenness centrality scores will be useful candidates in the actual lead user project? The relative lead userness scores alone cannot answer this question. Theoretically, users in this specific subgroup might have higher lead user characteristics compared with the mass of users in the forum, but they still might not be able to provide valuable input in the Airbus lead user project. Two analyses are offered in order to rule out this possibility.

First, the frequency with which novel solutions are developed in the two groups is compared. Lead user theory suggests that the likelihood of providing innovations should be higher among lead users than among average users (von Hippel, 1986). Lead userness and the development of novel solutions are highly correlated ($r = .483$, $p = .000$), and the correlation of betweenness centrality with the development of novel solutions is even higher ($r = .610$, $p = .000$). This provides clear evidence that it makes sense to identify the forum users with the highest levels of betweenness centrality in commercial lead user studies.

Second, solutions and solution-related knowledge from the posts provided by those 50 forum members as potential input to the Airbus lead user project are collected. We presented the solutions and knowledge to the executives of Airbus involved in the lead user project and asked them to evaluate the value of this input for the project. After careful inspection of the input, the executives stated that they would consider the input from the 50 forum members with the highest betweenness centrality to be highly valuable for the project. Therefore, one conclusion can be drawn, namely that identifying individuals with high levels of betweenness centrality in relevant online forums seems to be a viable shortcut for finding lead users in real-world settings.

General Discussion

In this paper, the social network position of lead users is examined as an approach for lead user identification. The three conducted studies, which use different empirical settings and different methods, yield the robust result that lead users have a distinctive social network position: they are positioned as bridges between different social groups. As social network positions of individuals can be mapped with modern Web mining tools quickly and at low cost, the results of these studies may show a new approach for the efficient and effective identification of lead users.

This paper has two major implications for lead user theory: First, it consolidates seminal conceptual work on lead users (von Hippel, 1986) and empirical research that focuses on the social network side of lead users (Kratzer and Lettl, 2008, 2009, 2011) with studies on the application of the lead user concept in innovation projects within companies (Herstatt and von Hippel, 1992; Lilien et al., 2002; Lüthje and Herstatt, 2004; Urban and von Hippel, 1988). The research shows that such a connection can indeed be established and thus leads to new insights: With the advent of the Internet and online social network applications, the social network perspective enables an effective and efficient identification of lead users. Second, the research extends and validates prior work on the social network perspective with respect to very different empirical settings and concrete lead user studies by companies in different sectors.

Currently a rapid growth in (online) user communities and a major shift of innovation activities from the offline to the online world is observed. The rise of the Web in general and online social network applications in particular has created unprecedented opportunities to identify lead users via social networks. Companies that do not take advantage of these opportunities may be left behind, as these online communities may very well develop into a major source of ideas and leading-edge innovation concepts (von Hippel, 2005).

From a managerial perspective, the insights from this study may help companies to increase the efficiency and effectiveness of lead user identification in different settings, including online social network applications. Based on the insights from this study, one market research design for lead user identification could be the collection and analysis of social network data from a user population within a certain domain using standard data mining software to examine online communication (as applied in Study 3).

This study is not without its limitations. This work is restricted in its conceptualization of simple and straightforward network measures. Limitations also arise because the studies only partly capture the inherent dynamics of social networks.

Future research could go beyond simple measures and utilize a highly diverse repertoire of social network analysis tools. Another avenue for future research would be longitudinal research designs in order to trace and follow individual changes in network configurations. Future research could also look deeper into the relationship

between a distinctive social network position and the likelihood that an individual will develop lead user characteristics. The paper argues that there is mutual causality between the social network position of an individual and the likelihood that the individual will display lead user characteristics. On the one hand, their need for innovation drives lead users to attain high levels of betweenness centrality. On the other hand, high betweenness centrality enables individuals to become lead users, as it provides them with leading-edge information on current trends. Although it may not be feasible to disentangle those two directions fully, it would be interesting to generate a better understanding of whether those two directions are at work simultaneously or sequentially, and whether one direction is more prevalent than the other. More research may also be necessary in order to understand which methods (and combinations of methods) of searching for lead users are most effective in a given situation, and how a betweenness centrality-based data mining software could be optimized for lead user identification purposes. Another future direction could be to combine the attributes of individuals and social networks into attribute networks. A successful combination would represent a major step toward finding a systematic method to identify lead users.

References

Barnowe, J. T. 1975. Leadership and performance outcomes in research organizations: The supervisor of scientists as a source of assistance. *Organizational Behavior and Human Performance* 14: 264–74.

Belz, F. M., and W. Baumbach. 2010. Netnography as a method of lead user identification. *Creativity and Innovation Management* 19 (3): 304–12.

Berg-Jensen, M., C. Hienerth, and C. Lettl. 2010. Forecasting the attractiveness of user-generated designs via online data. An empirical study within the LEGO user community. *Best Paper Proceedings of the Academy of Management Conference*, Montreal, Canada.

Bilgram, V., A. Brem, and K.-I. Voigt. 2008. User-centric innovations in new product development—Systematic identification of lead users harnessing interactive and collaborative online-tools. *International Journal of Innovation Management* 12: 419–58.

Borgatti, S. P. 1995. Centrality and AIDS. *CONNECTIONS* 18: 112–26.

Borgatti, S. P., M. G. Everett, and L. C. Freeman. 2002. *UCINET 6 for Windows*. Harvard, MA: Harvard, Analytic Technologies.

Brown, J. S., and P. Duguid. 2001. Knowledge and organization: A social-practice perspective. *Organization Science* 12: 198–213.

Burt, R. 1992. *Structural holes*. Cambridge, MA: Harvard University Press.

Cooke, M., and N. Buckley. 2008. Web 2.0, social networks and the future of market research. *International Journal of Market Research* 50: 267–92.

Cooper, R. 1981. An empirically derived new product project selection model. *IEEE Transaction on Engineering Management* 28: 54–61.

Dahl, W., and P. Moreau. 2002. The influence and value of analogical thinking during new product ideation. *Journal of Marketing Research* 39: 47–60.

Defares, P. B., G. N. Kema, E. Van Praag, and J. J. Van der Werft. 1971. *Syracuse-Amsterdam-Groningen Sociometrische Schaal*. Rijksuniversiteit Groningen: Subfaculteit der Pedagogische en Androgogische Wetenschappen.

Ebadi, Y. M., and J. M. Utterback. 1984. The effects of communication on technological innovation. *Management Science* 30: 572–885.

Fleming, L. 2007. Breakthroughs and the "long tail" of innovation. *MIT Sloan Management Review* 49: 69–74.

Franke, N., and S. Shah. 2003. How communities support innovative activities: An exploration of assistance and sharing among end-users. *Research Policy* 32 (1): 157–78.

Franke, N., E. von Hippel, and M. Schreier. 2006. Finding commercially attractive user innovation: A test of lead-user theory. *Journal of Product Innovation Management* 23: 301–15.

Freeman, L. C. 1979. Centrality in social networks conceptual clarifications. *Social Networks* 1: 215–39.

Frick, K., D. Guertler, and P. Gloor. 2013. Coolhunting for the world's thought leaders. *Proceedings 4th International Conference on Collaborative Innovation Networks COINs13*, Santiago de Chile. Available at: http://arxiv.org/pdf/1308.1160.pdf.

Füller, J., K. Matzler, and M. Hoppe. 2008. Brand community members as a source of innovation. *Journal of Product Innovation Management* 25: 608–19.

Gloor, P., P. J. Krauss, S. Nann, K. Fischbach, and D. Schoder. 2009. Web Science 2.0: Identifying trends through semantic social network analysis. *IEEE Conference on Social Computing*, Vancouver, Canada.

Guilford, J. P. 1950. Creativity. *The American Psychologist* 5: 444–54.

Hayes, A. F., and K. Krippendorff. 2007. Answering the call for a standard reliability measure for coding data. *Communication Measures and Methods* 1 (1): 77–89.

Healy, A. F., D. A. Havas, and J. T. Parkour. 2000. Comparing serial position effects in semantic and episodic memory using reconstruction of order tasks. *Journal of Memory and Language* 42: 147–67.

Heneman, H. G. 1974. Comparisons of self- and superior ratings of managerial performance. *The Journal of Applied Psychology* 59: 638–42.

Herstatt, C., and E. von Hippel. 1992. From experience: Developing new product concepts via the lead user method: A case study in a "low-tech" field. *Journal of Product Innovation Management* 9: 213–21.

Huston, L., and N. Sakkab. 2006. Connect and develop: Inside P&G's new model for innovation. *Harvard Business Review* 84 (3): 58–66.

Jeppesen, L. B., and K. Laursen. 2009. The role of lead users in knowledge sharing. *Research Policy* 38: 1582–89.

Kozinets, R. V. 2002. The field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research* 39: 61–72.

Kozinets, R. V. 2010. *Netnography: Doing ethnographic research online*. London: Sage.

Krackhardt, D. 1987. QAP partialling as a test of spuriousness. *Social Networks* 9: 171–86.

Kratzer, J., and C. Lettl. 2008. A social network perspective of lead users and creativity: An empirical study among children. *Creativity and Innovation Management* 17: 26–36.

Kratzer, J., and C. Lettl. 2009. Distinctive roles of lead users and opinion leaders in the social networks of schoolchildren. *The Journal of Consumer Research* 36: 646–59.

Kratzer, J., and C. Lettl. 2011. Die Identifizierung von Lead Usern über soziale Netzwerke: Eine empirische Untersuchung unter jungen Konsumenten. *Zeitschrift für Betriebswirtschaft* 81 (5): 83–110.

Krippendorff, K. 2004. *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.

Kunst, L., and J. Kratzer. 2007. Diffusion of innovations through social networks of children. *Young Consumers* 8: 36–51.

Leenders, R. T., J. M. Van Engelen, and J. Kratzer. 2007. Systematic design methods and the creative performance of new product units: Do they contradict or complement each other. *Journal of Product Innovation Management* 24: 166–79.

Lettl, C., C. Herstatt, and H. G. Gemünden. 2006. Users contributions to radical innovation. Evidence from four cases in the field of medical equipment technology. *R&D Management* 36: 251–72.

Lilien, G. L., P. D. Morrison, K. Searls, M. Sonnack, and E. von Hippel. 2002. Performance assessment of the lead user idea-generation process for new product development. *Management Science* 48: 1042–59.

Long, Y., and K. Siau. 2007. Social network structures in open source software development teams. *Journal of Database Management* 18 (2): 25–40.

Lüthje, C., and C. Herstatt. 2004. The lead user method: An outline of empirical findings and issues for future research. *R&D Management* 34: 553–68.

Lüthje, C., C. Herstatt, and E. von Hippel. 2005. User-innovators and "local" information: The case of mountain biking. *Research Policy* 34: 951–65.

Moriarty, R. T., and J. E. G. Bateson. 1982. Exploring complex decision making units: A new approach. *Journal of Marketing Research* 19: 182–91.

Morrison, P. D., J. H. Roberts, and D. F. Midgley. 2004. The nature of lead users and measurement of leading edge status. *Research Policy* 33: 351–62.

Morrison, P. D., J. H. Roberts, and E. von Hippel. 2000. Determinants of user innovation and innovation sharing in a local market. *Management Science* 46: 1513–27.

Mumford, M. D., and S. B. Gustafson. 1988. Creativity syndrome: Integration, application and innovation. *Psychological Bulletin* 103: 27–43.

Nelson, R., and S. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Belknap Press.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. *Applied linear statistical models* (4th ed.). Chicago, IL: Richard D. Irwin, Inc.

Neuendorf, K. A. 2002. *The content analysis guidebook*. Thousand Oaks: Sage Publications.

Nicholas, J. M. 1994. Concurrent engineering: Overcoming obstacles to unit work. *Production and Inventory Management Journal* 35: 234–46.

Nonaka, I. 1994. A dynamic theory of organizational knowledge creation. *Organization Science* 5 (1): 14–37.

Nyström, H. 1979. *Creativity and innovation*. New York: John Wiley & Sons.

Perry-Smith, J. E., and C. E. Shalley. 2003. The social side of creativity: A static and dynamic social network perspective. *Academy of Management Review* 28: 89–1073.

Polyani, M. 1966. *The tacit dimension*. Garden City, NY: Doubleday & Company.

Riggs, W., and E. von Hippel. 1994. Incentives to innovate and the sources of innovation: The case of scientific instruments. *Research Policy* 23 (4): 459–69.

Sawhney, M., G. Verona, and E. Prandelli. 2005. Collaborating to create: The internet as a platform for customer engagement in product innovation. *Journal of Interactive Marketing* 19: 4–17.

Schreier, M., S. Oberhauser, and R. Prügl. 2007. Lead users and the adoption and diffusion of new products: Insight from two extreme sport communities. *Marketing Letters* 18: 15–30.

Schumpeter, J. 1939. *Business cycles*. New York: McGraw-Hill Book Company, Inc.

Shalley, C. E. 1995. Effects of coaction, expected evaluation and goal setting on creativity and productivity. *Academy of Management Journal* 38: 483–503.

Sosa, M. E., S. E. Eppinger, and C. M. Rowles. 2004. The misalignment of product architecture and organizational structure in complex product development. *Management Science* 50 (12): 1674–89.

Spirtes, P., C. Glymour, and R. Scheines. 2000. *Causation, prediction, and search* (2nd ed.). Cambridge, MA: MIT Press.

Staber, U. 2004. Networking beyond organizational boundaries: The case of project organizations. *Creativity and Innovation Management* 13: 30–40.

Tietz, R., P. D. Morrison, C. Lüthje, and C. Herstatt. 2005. The process of user-innovation: A case study in a consumer goods setting. *International Journal of Product Development Management* 2 (4): 321–38.

Torrance, E. 1974. *The Torrance tests of creative thinking*. Bensenville: Personnel Press.

Urban, G., and E. von Hippel. 1988. Lead user analysis for the development of new industrial products. *Management Science* 34: 569–82.

Usher, A. P. 1954. *A history of mechanical invention*. Cambridge, MA: Harvard University Press.

von Hippel, E. 1986. Lead users: A source of novel product concepts. *Management Science* 32: 791–805.

von Hippel, E. 2005. *Democratizing innovation*. Cambridge, MA: MIT Press.

von Hippel, E. 2007. Horizontal innovation networks—By and for users. *Industrial and Corporate Change* 16: 1–23.

von Hippel, E., N. Franke, and R. Prügl. 2009. Pyramiding: Efficient search for rare subjects. *Research Policy* 38: 1397–406.

Vermunt, J. K., and J. Magidson. 2002. Latent class cluster analysis. In *Applied latent class analysis*, ed. J. A. Hagenaars and A. L. McCutcheon, 89–106. Cambridge, UK: Cambridge University Press.

Ward, T. B. 1994. Structured imagination: The role of category in exemplar generation. *Cognitive Psychology* 27: 1–40.

Wasserman, S., and K. Faust. 1994. *Social network analysis*. New York: Cambridge University Press.

Weimann, G. 1982. On the importance of marginality: One more step into the two-step flow of communication. *American Sociological Review* 47: 764–74.

West, M. A. 1990. The social psychology of innovation in groups. In *Innovation and creativity at work*, ed. M. A. West and J. L. Farr, 309–33. Chichester, UK: Wiley.

Wilensky, H. L. 1967. *Organizational intelligence*. New York: Basic Books.

Appendix A

Lead User Cluster Identification Process

After testing the performance of the lead user construct in Study 1, we conducted latent class analyses with all six items (Vermunt and Magidson, 2002). In order to determine the number of segments, we compared the models using the Bayesian Information Criterion until model fit stopped improving (Tables A1 and A2).

In order to assess the quality of the lead user identification process, we performed a number of *t*-tests. In order to execute these analyses, we formed two groups as follows: the participants in Study 1 were classified as "lead users" (cluster 1; $n = 34$) and "average users" (clusters 2–6; $n = 233$). In this way, we were able to check whether there were significant differences between lead users and average users in terms of lead user characteristics. Table A2 shows that the results were affirmative, as the lead user cluster shows significantly higher values for each of the items reflecting the two lead user characteristics of "trend position" and "expected benefit."

Table A1. Results of Latent Class Analysis—Study 1

Number of Segments	Bayesian Information Criterion	
1	5060.50	
2	4413.66	
3	4229.01	
4	4194.59	
5	4167.03	
6	**4103.55**	**= > Optimal cluster solution**
7	4198.96	
8	4212.17	
9	4225.75	
10	4345.61	

The latent class analysis for Study 1 resulted in a six-cluster solution.

Table A2. Results of *t*-Tests to Assess Quality of the Lead User Cluster Process—Study 1[a]

		Trend Position Statement 1		Trend Position Statement 2		Trend Position Statement 3		Expected Benefit Statement 1		Expected Benefit Statement 2		Expected Benefit Statement 3	
	n	Mean	*p*	Mean	*p*	Mean	*p*	Mean	*p*	Mean	*p*	Mean	*p*
Lead user	34	4.68	.000	4.91	.000	4.23	.000	4.81	.000	4.89	.000	4.73	.000
Average user	233	3.22		3.19		2.85		3.45		3.56		3.51	

[a] Two-tailed tests are reported; *t*-test.

Appendix B

Table B1. Coding Table for Lead User Characteristics—Study 3

Lead User Item	Typical Examples
Development of novel solution	Author: Edward429451 (CO) I know some of you guys are doing forced air & certainly boilers. Sometimes we get those intermittent shutdown problems and of course they won't shut down when we're there. Auto reset safeties fail under load, intermittent limit problems, relays, etc. It's cut & dried if they fail in front of you but if not what to do? TattleTales are single use devices that you clip on safeties in parallel that change color when the device opens, letting you see later what had opened and give you a clue. Trouble is they are single use and expensive and do you have any/enough on the truck when you need it? Make your own. I used two ¼″ F spade connectors, short pieces of wire and either alligator clips or two piggyback spade connectors (piggybacks for rollouts and limits that have no room to clip alligator clips on when there's no room without breaking the circuit) and heat shrink to pretty it all up. Then you just install an automotive type flat fuse, low amp and tie it on. When the safety opens the power outside goes to ground and pops the fuse, allowing you to later see what opened. It will still allow the unit module/lockout to be reset so they will still be able to get heat (intermittently) cause its in parallel until you return. Reusable, just replace the fuse. On relays you can just put in an appropriate larger fuse. Big time saver, don't have to sit there cycling the unit for hour(s) hoping it will shut down and it wont. BTDT. Post your own McGyver type tricks of the trade. . .
Providing solution-related knowledge	Author: dlh (TX) The first thing I would do is look for a steel nipple at the water heater inlet and outlet. Being the poster has galvanized piping it is a safe bet that one or both connections at the heater are occluded.
Providing product-related knowledge	Author: redwood (CT) I didn't say don't use PEX! I said don't use Zurn fittings. Watts would be a good choice for you. Raven Products makes a SSC crimping tool that many supply houses sell for about $50.
Providing advice to forum peers	Author: redwood (CT) The water might have been leaking out from the tank faster than it could get out from under the jacket. So at some point the jacket blew out at the seam. I've seen it a few times now with the foam insulated water heaters. That foam is tight and bonds pretty good.
Making references to state-of-the-art/new products or trends	Author: waukeshaplumbing (WI) AAV's are going to be illegal here in WI in 2011 . . . the trend is away from them . . . why install something which is going to fail . . . and then installing them everywhere . . . i wouldnt do it. Hire a plumber
Identification of a problem/ articulating a problem	Author: e-plumber (NY) I have to agree. I equate these types of fittings with compression unions, fittings, etc., which I use ONLY for a temporary "quick fix" in certain situations to get things up and running. Maybe it's me but I don't trust them, I would feel 100% confident with a solder joint over any type of "push-fit" fitting. My opinion may also be because of the part of the country that we live in where copper piping w/ sweat joints have been used for the past 50+ years for residential domestic & heating lines, most all inspectors here would frown on anything else. e-plumber
Articulating dissatisfaction/ disappointment with current products	Author: mjb1962853 (NY) I'm NOT a plumber, but I'm sure someone will still yell at me because what I'm about to suggest will surely cause the universe to implode. Many electric heaters have detents or marks at 125 deg F and I can tolerate that during the summer months. During the cold winters here, we need to set ours at 135 deg F (verified with a separate thermometer). We all know this increases the potential for a scalding accident, but everyone in this house understands the risk and has NO problem with it.