The frequency of end-user innovation: A re-estimation of extant findings

Nikolaus Franke*, Florian Schirg, Kathrin Reinsberger

Institute for Entrepreneurship and Innovation, Vienna University of Economics and Business (WU Vienna), Welthandelsplatz 1, 1020 Vienna, Austria

ABSTRACT

Recent studies have found that large numbers of consumers innovate. In our study, we provide a re-estimation of the figures provided in the extant literature. We do so by conducting a study in which we apply two different methods of data collection: (1) telephone interviews, the method considered most valid in previous research, and (2) personal interviews, which involve much higher effort but induce better individual recollection. Using telephone interviews, we measured a user-innovation frequency of 10.6% in our sample. In stark contrast, personal follow-up interviews resulted in a frequency of 39.7%, indicating a considerable underestimation in extant research. We then used the correction factor generated to re-estimate findings on user innovation frequency in Finland, Japan, Korea, Sweden, the UK, and the USA. It appears that user innovation is indeed a mass phenomenon that should not be overlooked by policymakers or firms.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction: recent studies on user innovation frequency

The dominant view on innovations has been that they “originate from producers and [are] supplied to consumers via goods and services that [are] for sale” (Baldwin and von Hippel 2011, p. 1399). Since the 1970s and 1980s, a number of studies have challenged this view, finding that many of the most important innovations originate from users, not from producers (see Von Hippel, 1988, 2005 for overviews). Recently, this line of research has been enriched by complementary studies that analyze the frequency of user innovators among representative samples of end-users or consumers. A seminal study carried out in the UK found that 6.1% of all adult consumers had created new products or modified existing products in an innovative manner in the three years prior to being interviewed (Von Hippel et al., 2012). A series of studies in other countries have confirmed that user innovation is quite a widespread activity (Finland: 5.4%, De Jong et al., 2015; Japan: 3.7%, Ogawa and Ponttanalert, 2013; Korea: 1.5%, Kim, 2015; Sweden: 7.3%, Bengtsson, 2015; USA: 5.2%, Ogawa and Ponttanalert, 2011).

Total user investments in innovation are even estimated to be higher in value than domestic R&D spending by industry (Von Hippel et al., 2012), and the efficiency of such investments is likewise higher (Hienith et al., 2014). All this supports the idea that we are in the midst of a paradigm shift from traditional producer innovation to user innovation (Baldwin and Von Hippel, 2011). Documenting the prevalence of user innovation is important because figures on this phenomenon are largely excluded from official statistics.

In this study, we argue that the frequency of user innovation reported in recent representative studies has been underestimated for methodological reasons, in particular due to the data collection method used. We investigated the extent of underestimation in an empirical study by collecting data on user innovation frequency using telephone interviews, the data collection method considered most valid in extant research. We then investigated how many “non-innovating” consumers (i.e., those who claimed that they had never innovated) actually had innovated but simply did not recall their activities during the telephone interview. In extensive face-to-face interviews, we found that the proportion of consumer innovators is actually almost four times as high as the one indicated by telephone interviews. On this basis, we re-estimated user innovation frequencies from extant research. The resulting figures are likely to be far closer to the true proportions. Finally, we close with a brief discussion of our results and the limitations of our study.

* Corresponding author.
E-mail addresses: nikolaus.franke@wu.ac.at (N. Franke), florian.schirg@wu.ac.at (F. Schirg), kathrin.reinsberger@wu.ac.at (K. Reinsberger).

1 http://www.wu.ac.at/entrep.
2 User Innovation Research Initiative: https://www.wu.ac.at/entrep/forschung/userinnovation.
2. Theory: why extant studies on user innovation frequency might underestimate the prevalence of innovation activity

The data collection method considered most valid in extant research on user innovation frequency is computer-aided telephone interviews (Bernard, 2002; Carr and Worth, 2001). This “gold standard” was used in the seminal study by Von Hippel et al. (2012) as well as in a very recent study by Bengtsson (2015). In these interviews, trained interviewers asked respondents to give details on their innovation activities in the previous three years. In order to help interviewees recall the occurrence of such behavior, interviewers used certain cues. For example, without asking directly about “innovation” activities, they asked about the subjects’ favorite hobbies and whether they had ever “created” a new product or “modified” an existing product in a novel way. This interview technique involves a large investment of time and money compared to written (online or paper-and-pencil) questionnaires. The main benefit is that the technique provides a much richer and more detailed depiction of user innovation activities than simple questionnaires. However, notwithstanding the considerable effort involved, this method can still lead to an underestimation of user innovation activity.

It is well known from research into survey methodology that self-reporting of past behavior is not perfectly accurate (Tourangeau et al., 2000). When asked about past activity, people cannot scan their memory like a computer that scans a storage medium for a given search string and reveals all hits immediately. When we ask people about past behavior, a number of problems might skew their responses (Schwarz and Oyserman, 2001; Sudman et al., 1986): People may (1) misunderstand the question, (2) fail to recall the relevant behavior, (3) apply biased inference and estimation strategies in arriving at an answer, (4) be unable to map the answer onto the response format, and (5) intentionally “edit” or revise their answers because of social desirability, self-presentation concerns, and other reasons. Depending on the nature of the question, on the situation, and on the characteristics of the interviewee, these problems may result in under-reporting or over-reporting of the behavior in question (Belson, 1981).

In the studies by Von Hippel et al. (2012) and Bengtsson (2015), trained professional interviewers conducted telephone interviews and provided clarification when necessary. This meant that the potential for misunderstanding (Problem 1) and the incorrect mapping of responses (Problem 4) were probably not very great. The problem of inference and estimation bias (Problem 3) could also largely be ruled out, as the objective of the studies was to measure the mere occurrence of user innovation (yes/no), not its frequency at the individual level. However, the problems relating to recall and editing (Problems 2 and 5) are likely to have had an impact here.

The problem of revising or editing answers (Problem 5) brings about potentially opposing effects. On the one hand, the qualitative, more extensive, dialogue-oriented interview technique may have induced interviewees to over-report (i.e. to invent or exaggerate) their user innovation behavior due to social desirability (Arnold and Feldman, 1981; Kreuter et al., 2008). Respondents might have quickly figured out that their own innovation activities, and they may have wanted to make a favorable impression on the interviewer. There is evidence that such problems occur. Von Hippel et al. (2012), for example, state that 199 respondents (17% of their total sample) reported innovations that did not match the authors’ original definition. Fortunately, researchers have been able to correct these false positives by examining the details of the innovations revealed.

On the other hand, under-reporting of innovation activity may also be a problem. Some interviewees may have concealed their innovations out of fear that their intellectual property could be exploited (Franke et al., 2013). Compared to face-to-face interviews, telephone interviews are sometimes criticized because they may create difficulties in developing rapport and trust (e.g. Sturges and Hanrahan, 2004). It is difficult to assess the extent to which such non-disclosure affects the figures reported in consumer innovation studies. Generally, user innovators are found to reveal their innovations freely because they have no commercial interest in exploiting them (Harhoff et al., 2003). This would indicate that non-disclosure is, in fact, unlikely to play a major role in extant studies on user innovation frequency. However, we must keep in mind that we know only little about the proportion of user innovators who do not reveal their innovations, as they may easily remain unidentified for this very reason. The phenomenon of user entrepreneurship (Shah and Tripsas, 2007), i.e. user innovators starting companies on the basis of an innovation they originally developed for themselves, allows us to speculate that keeping an innovation secret might constitute an incentive for at least some user innovators. It is well known that entrepreneurs reveal information about their business ideas only to trusted institutions (Bergh et al., 2011). Moreover, user innovations in fields that are potentially embarrassing or illegal, e.g. questions of personal hygiene, sexuality, drugs, weapons, etc., will only be revealed once sufficient trust or rapport has been established between the interviewer and interviewee. At the very least, this means that as far as telephone interviews are concerned, we simply cannot rule out the possibility that answer editing might have led to an underestimation of the true proportion of user innovators.

The most severe underestimation problem, however, may result from forgetting (Problem 2). It is normal for memories to fade away; in other words, information traces decay, even when the event is subjectively interpreted as significant. One humorous example can be found in the German version of the “Who wants to be a millionaire?” TV show in 2013, when the well-known Bavarian soccer journalist Waldemar Hartmann was used as the “Ask an expert” telephone lifeline. When asked which of the four teams shown had never won the soccer world championship as the host country, he completely forgot Germany’s triumph in the final in Munich (Bavaria’s capital!) in 1974 and answered — to the spectators’ amazement: “There is surely only one. Germany has, of course, not won a world championship as a host country . . . not yet in their own country. You can read this in my book . . . “. We have all probably experienced such unbelievable lapses ourselves. More systematic evidence is provided in the classic study by Cannell et al. (1965), who found that 97% of their respondents recalled an episode of hospitalization when interviewed ten weeks after the event, but only 68% did so when interviewed a year later. In consumer innovation studies, the time horizon to which the interview refers is the previous three years. Additionally, we cannot even be sure that the respondents considered the innovation activity as significant and worth keeping in mind (personal accounts by the authors of this article suggest that some user innovators regard their endeavors as something that simply “goes without saying”). Moreover, “creating new products or modifying existing products in an innovative manner” is not a common category in an individual person’s memory (Belli, 1998). As Schwarz and Oyserman (2001, p. 137) put it:

“The structure of autobiographical memory can be thought of as a hierarchical network that includes extended periods (such as ‘the years I lived in New York’) at the highest level of the hierarchy. Nested within these high-order periods are lower-level extended events pertaining to this time, such as ‘my first job’ or ‘the time I was married to Lucy.’ Further down the hierarchy, are summarized events, which correspond to the knowledge-like representations of repeated behaviors noted above (e.g. ‘during that time, my spouse and I quarreled a lot’). Specific events, such as a particular instantiation of a disagreement, are represented...
at the lowest level of the hierarchy. To be represented at this level of specificity, however, the event has to be rather unusual."

Retrieving past behavior from memory thus involves pursuing numerous top-down pathways across life themes that combine extended events, as well as moving in parallel across life themes that involve contemporary and consecutive events (Belli, 1998). This can mean that trace information exists but is (temporarily) inaccessible because it is not triggered by the right associative memory cues that were present when the memory was encoded, such as environment, situation, mood, state of mind, etc. (Neath, 1998). This form of forgetting is known as "cue-dependent forgetting" (Tulving, 1974). Thus, while user innovators may well have stored the information on how they fixed a problem or came up with an innovative way of doing something, the cues used in the interview situation may not have been effective in aiding the retrieval of this information. Generally, searches within one’s memory take considerable time (Reiser, Black and Abelson 1985), and their outcome depends on the usefulness and number of network entry points (cues) provided (Schwarz and Oyserman, 2001). Two aspects of the telephone interview permit us to surmise that not all respondents remembered instances of innovation activities: First, the inevitable pauses and moments of silence necessary for extended memory search during telephone interviews will typically make respondents feel uneasy, and they may be tempted to abandon the process too early, resulting in a failure to retrieve information (Schwarz and Oyserman, 2001). Second, telephone interviews are limited to verbal cues (Carr and Worth, 2001) and are therefore not as rich as personal interviews, in which (more effective) visual cues can also be given (McCoyd and Kerson, 2006). In sum, there are several reasons to suspect that extant studies on user innovation have systematically underestimated its prevalence.

3. Method: telephone interviews followed by personal interviews

Our study is based on a two-stage design: In an initial step, we screened for consumer innovators using telephone interviews. This approach simply emulates the most valid modus operandi of extant studies in this line of research. In a second step, in order to identify further user innovators, we conducted extensive face-to-face interviews with those respondents who had not previously identified themselves as user innovators during the telephone interview. Personal interviews are generally assumed to be superior to telephone interviews in several ways (Novick, 2008). For example, when conducted with due care, they facilitate the generation of trust more effectively (Creswell, 1998), they entail less pressure for an immediate response (Carr and Worth, 2001), and they allow the use of visual cues (Garbett and McCormack, 2001). All of these factors are likely to help improve respondent recall (Baddeley, 1990).

For the first step, we randomly selected 500 persons from the Vienna telephone directory. A comprehensive list of Vienna’s inhabitants was not available, so we used the directory as a proxy. Prior to our phone calls, we sent an invitation letter by post to the chosen addresses to underline the serious nature of our research. Seventy entries were neutral losses (e.g. unknown address, moved away, etc.), leaving us with a sample of 430 potential respondents for telephone interviews. Among them, 137 remained unreachable even after three attempts, and a total of 113 declined an interview. At that point, they were unaware of the specific topic and thus exhibited unsystematic non-response behavior. In the end, 180 persons agreed to participate in the telephone survey. Fourteen of them had to be excluded ex post (for neutral reasons), leaving us with a sample size of 166. We used the same interview script as von Hippel et al. (2012; obtained by the authors) and applied the same identification and validation procedure as Bengtsson (2015), De Jong et al. (2015), Kim (2015), Ogawa and Pongtanalert (2011), Ogawa and Pongtanalert (2013), and Von Hippel et al. (2012). In our case, 14 of the individuals who reported innovative activities had to be treated as false positives, as their activities did not match the accepted definition of an innovation. We then asked the 134 respondents who stated that they had not innovated in the past whether they would agree to be questioned in a face-to-face interview “on other topics”. A total of 34 agreed to do so (26%, with no indication of a response bias).

In order to create a positive atmosphere, i.e. one suitable for the development of trust and rapport, we made efforts to conduct the interviews at the participants’ homes. The interviews lasted between 25 and 90 min (average: 35 min). In the interviews, we stimulated the participants’ memories of innovation activities by showing and explaining visual examples of various consumer innovations, depicted on two large posters. Visual examples are effective cues for retrieving information because they can trigger specific associations (Smith and Vela, 2001; Wagenaar, 1986, 1988). Our impression was that this worked very well, and both the atmosphere and cues helped interviewees retrieve information about innovation activities that had probably been inaccessible during the telephone interviews. As in the telephone interviews, we asked respondents for more details whenever they mentioned an innovation. After the interviews, trained evaluators (including the second author of this study) examined the innovation accounts extensively. The same definition of user innovation and the same control procedure was used as in Von Hippel et al., (2012). This cross-check was done in order to increase the (internal) validity of the final results. Of the innovators, eight were excluded as false positives in this second step.

4. Results: user innovation is far more prevalent than telephone interviews suggest

We find that the true number of consumer innovators is substantially higher than the figures derived from telephone interviews. The first column in Table 1 shows the frequency of user innovators as measured by telephone interviews. The percentage we found in our sample (10.8%) is actually somewhat higher than that found in other extant studies. This may be attributed to the historically high level of creativity in Austria (Grassl and Smith, 1986) and particularly in Vienna (Janik, 2000), or simply to the fact that our sample was restricted to urban dwellers; innovation activity is known to be stimulated by agglomeration and geographical proximity (Audretsch, 1998). The figures for amounts invested, IP protection, information sharing, and adoption by others correspond nicely to the patterns found in the other studies.

The second column reports the innovation activities of those 34 interviewees who had not come up with an innovation in the telephone interviews. We find that during the face-to-face interview, almost a third of them – 11 individuals, or 32.4% – remembered that they had in fact innovated and proceeded to give detailed accounts of the processes and the results. They represent those user innovators who remain unnoticed in studies that do not use personal interviews. Their innovations appear somewhat less important in terms of investment, IP protection effort, and diffusion than those reported in the telephone interviews, suggesting that “innovation importance” probably also plays a role in the process of recall. Nonetheless, the additional innovations mentioned met the definitional criteria set by Von Hippel et al. (2012) and subsequent studies, and the most important innovations clearly exceed the mean level of significance of those found via telephone interviews.
5. Discussion: re-estimation of extant figures and suggestions for further research

First, our findings enable us to re-estimate extant findings on user innovator frequency (Table 2). We do so by employing the correction factor derived from our study (i.e. 3.7).\(^2\) Note that we also use this factor for the studies based on online questionnaires, a data collection method which is known to provide more downward-biased figures than telephone interviews when it comes to retrieving past behavior from memory (Tourangeau et al., 2000; Harris and Brown, 2010; Phellas et al., 2012). The participants’ inability or unwillingness to provide written texts, the impossibility of asking for clarification, and the difficulties of understanding and interpreting answers (Bryman and Bell, 2015) may also lead to an underestimation of user innovation. After all, it is necessary to check whether a self-reported “innovation” actually is sufficiently novel. All this makes our re-estimation of the figures provided by De Jong et al. (2015), Kim (2015), Ogawa and Pongtanalert (2011) and Ogawa and Pongtanalert (2013) particularly conservative.\(^3\)

The resulting figures imply that in these six countries alone, a total of 100 million consumers have innovated in the past three years. This reinforces a key implication cited in the original studies, namely that user innovation is a mass phenomenon.

This not only means that official statistics need to be complemented; it also indicates that we need to become much more aware of users as generators of new products and services. It also implies that specific policies should be adopted in order to nurture and support this source of innovation in a systematic manner. After all, consumer innovation increases welfare (Gambardella et al., 2014).

For companies, it means that consumers are an even richer source of innovation ideas than previously believed. It would therefore make sense to exploit this phenomenon, e.g. via lead user studies, toolkits for user innovation and design, and crowdsourcing innovation contests.

Second, we certainly cannot claim to have identified a nomological correction factor, a constant that is independent of time, place, and situation. Assuming that it is too expensive to collect data only by personal interviews in future research on user innovation frequency, we suggest that the following procedure be used to estimate true innovation frequency and the respective correction factors:

(1) Obtain data by telephone or online survey; screen the innovators identified in the original sample for false positives.

(2) Draw a random sub-sample of at least \(n = 30\) from those consumers who reported not having innovated or reported false positives.

(3) Train interviewers, interview sub-sample in their homes, create a trustful and relaxed atmosphere, and use visual cues to depict user innovations.

(4) Screen innovators identified for false positives.

(5) Estimate the true innovation frequency of the original sample as follows:

\[
\text{Frequency}_{\text{true}} = \left[ \frac{I_o + I_f + (S_o - I_o - S_f) \times I_f}{S_f} \right]
\]

\(I_o = \) innovators identified through original method in original sample, \(I_f = \) innovators identified through face-to-face interviews in sub-sample, \(S_o = \) original sample, \(S_f = \) sub-sample

(6) Estimate correction factor as follows:

\[
\text{Correction factor} = \frac{\text{Frequency}_{\text{true}} \times S_o}{I_o}
\]

Such a procedure would be particularly important where conditions differ systematically from our study. For example, scholars might also attempt to estimate the frequency of service innovations by users, a study that we believe is long overdue. Here, the correction factor would probably be much higher than in our study, as it might be even more difficult to retrieve information on an “abstract” entity such as a self-innovated service (Oliveira and von Hippel, 2010; Schulteß et al., 2010).

6. Limitations

Our study is based on a number of assumptions, all of which might be criticized for being simplifications. Regarding the estimation of the true number of user innovators in our Vienna sample, we first assume that there are no self-selection effects, neither in the...
original sample nor in the sub-sample. However, we cannot rule out the possibility that consumers who had a vague feeling that they had innovated in some way were also the ones who agreed to participate in the personal interviews (similar to a “tip of the tongue” effect, Brown, 1991). This would mean that the sub-sample is not random and that our correction factor might be inflated. Second, we assume that our sub-sample is large enough to provide a robust basis for calculating a correction factor. In fact, n = 34 is close to the minimum acceptable level for such an endeavor. However, a larger sample would increase reliability. Third, we also assume that personal interviews allow us to capture the true number of user innovators. This might be wrong. There probably are people in our sub-sample who have innovated but are still unable or unwilling to reveal those activities for one of the reasons mentioned above. This would suggest that our correction factor is too low. In short, the correction factor needs to be used with considerable care. However, despite all these drawbacks, it certainly improves attempts to estimate the true frequency of user innovators. Fourth, strictly speaking, a sample drawn from the local telephone directory is not representative of the overall population (Sudman, 1973; Smith et al., 1997). However, this method is a standard approach in population statistics (Sekaran and Bougie, 2009; Cramer and Howitt, 2004) and we are primarily interested in estimating a correction factor, not user innovation frequency as such. Fifth, we drew the sub-sample only from those consumers who indicated that they had not innovated, and we ignored the false positives. This creates a slight conservative bias in our correction factor. We find it plausible that the percentage of true user innovators is higher among consumers who reported some creative achievement than among those who did not (in the telephone interview). Sixth, our derivation of the correction factor on the basis of the Vienna sample assumes that the specific time, place and context are irrelevant. Again, this is a bold assumption. The time of year, current events, cultural backgrounds, the interviewer’s capabilities, etc. may all influence the results. These limitations underscore our suggestion that future studies on user innovation frequency among consumers should correct their estimates using self-generated correction factors. A final caveat of our study relates to our conclusion that end-user innovation frequency has been vastly underestimated and thus deserves far more attention. Might data sets on firm innovation (e.g. the Community Innovation Survey) suffer from similar limitations concerning data collection? Indeed, a recent study suggests that much innovation by and within firms is “informal” and may therefore not even be known to the person being interviewed (Hartmann and Hartmann, 2015). Could this mean that the relative overall importance of user innovation may remain unchanged? The honest answer is that, as yet, we simply do not know. As with the other limitations, more research is certainly warranted.

Acknowledgements

The authors wish to thank Aneta Bryla and Andreas Kardinal for their help in collecting the data and Chris Anderson for proof-reading the article. In addition, we are grateful for the valuable comments provided by three anonymous referees, who helped us improve this paper significantly. An earlier version of the paper was presented at the 12th International Open and User Innovation Workshop 2012 in Cambridge, MA.

References

Kim, Youngbae, 2015. Consumer user innovation in Korea: an international
Kreuter, F., Presser, S., Tourangeau, R., 2008. Social desirability bias in CATI, IVR,
Novick, Gina, 2008. Is there a bias against telephone interviews in qualitative
Schulteß, P., Wegener, S., Neus, A., Satzger, G., 2010. Innovating for and with your
service customers: an assessment of the current practice of collaborative
service innovation in Germany. Procedia Soc. Behav. Sci. 2, 6503–6515.
Schwarz, Norbert, Oyerman, Daphna, 2001. Asking questions about behavior:
Approach, 5th ed. Wiley and Sons, United Kingdom.
Shah, S.K., Tripsas, M., 2007. The accidental entrepreneur: the emergent and
collective process of user entrepreneurship. Strateg. Entrepreneur. J. 1 (1–2),
123–140.
Smith, W., Mitchell, P., Attebo, K., Leeder, S., 1997. Selection bias from sampling
frames: telephone directory and electoral roll compared with door–to–door
population census: results from the Blue Mountains Eye Study. Aust. N. Z. J.
Public Health 21 (2), 115–236.
Sudman, S., Bradburn, N.M., Schwarz, N., 1996. Thinking About Answers: The
Application of Cognitive Processes to Survey Methodology. Jossey–Bass, San
Francisco, CA.
Res. 10 (2), 204–207.
Tulving, Endel, 1974. Cue–dependent forgetting: when we forget something we
once knew, it does not necessarily mean that the memory trace has been lost;
it may only be inaccessible. Am. Sci., 74–82.
sector innovation in consumer products: findings from a representative study
in the United Kingdom. Manage. Sci. 58 (9), 1669–1681.
years. Cognit. Psychol. 18, 225–252.
and retrieval from autobiographical memory. In: Gruneberg, M.M., Morris, P.E.,
Sykes, R.N. (Eds.), Practical Aspects of Memory: Current Research and Issues 1,