The Quality of Online Registration Information

Factors influencing user decisions to reveal authentic personal information to online marketers as part of a perceived barter.

Andreas Neus

IBM e-business Innovation Center
Alter Wandrahm 10
20457 Hamburg, Germany
andreas.neus@de.ibm.com

*It ain't so much the things we don't know that get us into trouble.
It's the things we know that ain't so.*
- Artemus Ward
(Huff 1954)

Abstract

The online-collection of user information is becoming more and more popular due to its favorable economics in comparison with traditional methods of data collection. The practice of "online-registration" allows companies to force potential customers to fill out arbitrary forms as a prerequisite to granting some incentive. However, consumers are increasingly concerned about their online privacy. This practice of online-registration being perceived as a barter rather than a voluntary disclosure, therefore the quality of the information thus gathered is questionable. This paper presents the key results of a study on the factors influencing the disclosure of personal information as part of an online registration. More than Three Hundred Internet users were interviewed in Germany with a traditional paper-and-pencil survey to examine the decision process leading to the disclosure of either authentic or falsified personal information.

1 Introduction

Interactive marketing, or creating a dialogue with the customer, with the help of databases that save all the information received from the client, provides a more focussed and precise target, at the same time leading to a more satisfied and loyal customer who takes the initiative. The challenge is to create that possibility and then to benefit from it.
IMO – Information Market Observatory, European Commission (1995)

215

Against the background of what has been called "Database Marketing" and has more recently expanded into "One-to-One Marketing", the Internet has given rise to a new kind of information collection, with some properties that are quite unlike anything that has been done offline, if only in scope. The so-called "online registration" is the process whereby a company offers some kind of incentive online (information, software, service, sweepstakes etc.) in return for filling out a registration form with more or less sensitive personal information – usually at the very least an e-mail address. Traditional sweepstakes and mail-in product-registration cards, which both promise (a chance for) a reward can be seen as the paper-based ancestors of today's online registration.

However, given the interactive medium in which online registrations occur, the incentives offered in exchange for personal information can be promised and delivered much more credibly than the – somewhat more abstract – chances of winning a sweepstakes or getting "better service and extended warranty" promised in the offline world.

The downright *barter* of personal user information in exchange for a concrete incentive such as access to information, software or services has become a reality online. In contrast to a postcard, the Internet protocols even allow a company to *force* the user to fill out the answer fields to some or all of the questions.

2 Barter and forced compliance

These two aspects of online registration – barter and forced compliance – set the process apart from the more "voluntary" act of filling out traditional registration cards. Due to the proliferation of online registrations, users are faced with a trade-off decision between the *cost* of revealing personal information and the the utility and attractiveness of an *incentive* almost on a daily basis.

Little has been published to date about this process and the factors influencing this trade-off decision regarding online-registrations, except that there are indications of a significant percentage of intentionally false information given during online registrations. I.e. in the 9th GVU Survey (GRTC, 1998), more than 50% of those polled stated that they had provided false information on a web-site brfore. Another indication of the low quality of gathered information can be seen in the high percentage of invalid e-mail addresses collected by online-registration found in some studies (Gräf & Heidingsfelder, 1999).

While functional e-mail addresses can be easily enforced by sending a required password or even embedding the e-mail-address in the software (as practiced by GameSpy Industries, 1999), other information is not as easily verified.

3 An interesting market

People who are willing to take the time to fill out a registration form online (whether truthfully or not), in order to receive some incentive, are a very interesting market segment: They are a

self-selected sample that has somehow come across the site in question and was interested enough in the "free" product or service the site offered to spend a few minutes of precious time filling out a form. These are wonderful prospects for a "full" or "pro" version of the offering – much more precise than any traditional demographic target segment.

The problems that traditional marketing campaigns incur by targeting a segment of which only a tiny fraction happens to be seriously interested in an offering, do not apply. The problems here are based not on the *sample* quality, but on the quality of the *information* gathered through the online-registration, because everything that is "known" about the sample may be wrong. Therefore, the challenge to marketers is to understand which factors influence the likelyhood of this highly interesting and focussed sample of prospects, providing more accurate and useful information.

4 The Survey

In order to examine the factors influencing the decision to provide accurate personal infomation in the context of an online registration, a paper-and-pencil survey was developed at the Institute of Psychology's Department for Research Methodology and IT, Bonn University in Germany.

An on-line questionnaire, while certainly more economical, was not a viable alternative, as it would have resulted in a recursive contamination – with the object of observation (online-questionnaires) also being the instrument of observation.

Along similar methodological considerations, a cluster-sampling approach was used to recruit the 352 subjects for this survey. Whereas, in traditional surveys, a sample made up of respondents of ramdomly solicited individuals is often preferable, the typically low response rates of i.e. 5-20% would be prohibitive for a study specifically examining factors influencing the willingness to provide information in response to a survey, as this would again result in a contamination of instrument and object of measurement.

Instead, clusters of defined groups (such as faculty and advanced student classes at Bonn and Aachen University, and several small IT companies' employees) were selected for the sample to ensure a high compliance and return rate, in order to minimize the non-response bias. With a 92% return rate and a total of 352 thoroughly answered questionnaires, this goal can be said to have been achieved.

The predominantly student sample cannot, of course, be said to be representative of the German online population. Interestingly, however, examination of the non-student part of the sample (about 20%) failed to reveal systematically different behaviour patterns from the student population.

5 Survey Results

When evaluating the results of the survey, it is important to note they represent statements of Internet users regarding their own behaviour and predispositions, which may – or may not – coincide with their actual behaviour, as people are prone to a number of biases when self-reporting (i.e. social desirability bias, availability heuristic, etc.).

But as the ambitious goal of the survey was to examine the factors influencing the decision process underlying the overt behaviour with regard to its impact on information quality, a purely behavioristic approach would not have been feasible.

In the scope of this paper, we are focusing on the result of the survey in respect to five key questions:[1]

1. Which factors increase the likelihood of revealing accurate personal information?
2. How much faith do users have in web-sites' compliance with their privacy policies?
3. Which are the preferred strategies for providing information in order to gain access to an incentive?
4. What are the strategies for dealing with requests for information the user does not wish to provide?
5. Which kinds of information are users prepared to reveal to someone they are doing business with online?

5.1 Which factors increase the likelihood of revealing accurate personal information?

The subjects were asked to mark for each of 11 factors, determined in a pre-study to be related to the likelihood of providing information to a web-site, whether or not these factors could have a positive influence on their decision to provide accurate personal information to a web-site. (Figure 1)

Question
"What could have a positive influence on your decision to reveal accurate personal information about yourself on a web-site?"

Key result:
The five most important factors are all privacy-related. Classic Internet incentives such as access to software, information or sweepstakes came out at the bottom with regard to the revelation of *accurate* personal information.

---

[1] A more detailed discussion will be published in a forthcoming German-language book by publisher PACE Bonn, Germany
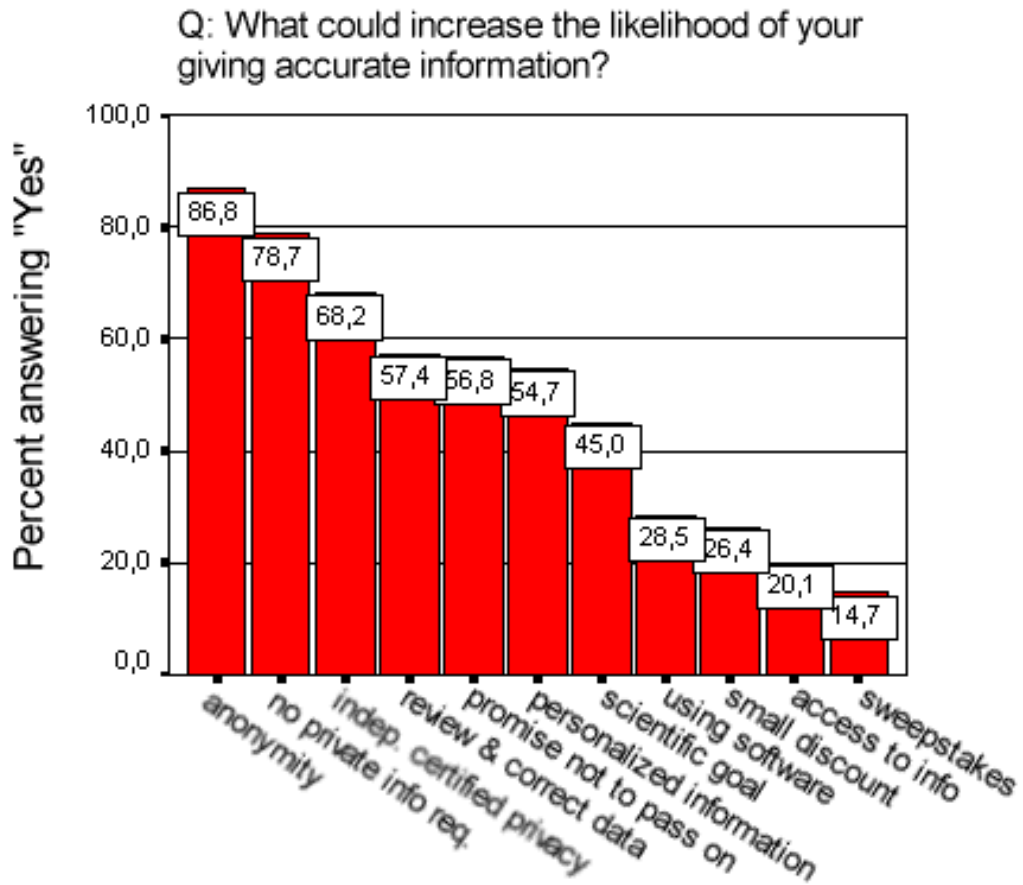
Figure 1 Factors that increase the likelihood of providing accurate personal information to a web-site.

The factors were (in order of importance):

|		
1)	anonymity (no identifying information is collected)	86.8 %
2)	no "private" information is required	78.7 %
3)	an independent entity certifies the site's privacy compliance	68.2 %
4)	the chance to review and correct the stored information	57.4 %
5)	the promise of not passing the information on to third parties	56.8 %
6)	a personalized information service (dates, news of interest)	54.7 %
7)	the scientific (vs. commercial) goal of the data collection	45.0 %
8)	answering the questions is a prerequisite to using a software	28.5 %
9)	answering the questions is a prerequisite to a small discount	26.4 %
10)	answering the questions is a prerequisite to accessing information	20.1 %
11)	answering the questions is a prerequisite to taking part in sweepstakes	14.7 %

The top factor, being cited as relevant by 87% of the subjects, is the condition of anonymity, followed by four more factors clustered around the concept of online privacy and control over one's personal information. These could be characterized as *Hygiene Factors* rather than true

Motivators (compare Herzberg's "Two Factor Theory", i.e. Herzberg et al., 1959). The only real *incentive* that could sway more than half of the subjects into providing accurate information was a personalized information service. However, the classic incentives for registration used by most web-sites (access to some information or software) come out last with regard to the revelation of *accurate* personal information. This supports the anecdotal evidence that online registration databases suffer from low quality information with pseudonyms such as "Fox Moulder, FBI" and "president@whitehouse.gov" apparently being particularly popular.

One troublesome result is that this also calls into serious question the quality of many of the online-surveys (i.e. W3B) which try to attract subjects by either offering sweepstakes or access to detailed result information of that survey.

5.2 How much faith do users have in web-sites' compliance with their privacy policies?

For this question, users were asked to rate their agreement or disagreement on a 4-item Likert-Scale: agree fully, agree somewhat, disagree somewhat, disagree fully. (Figure 2)
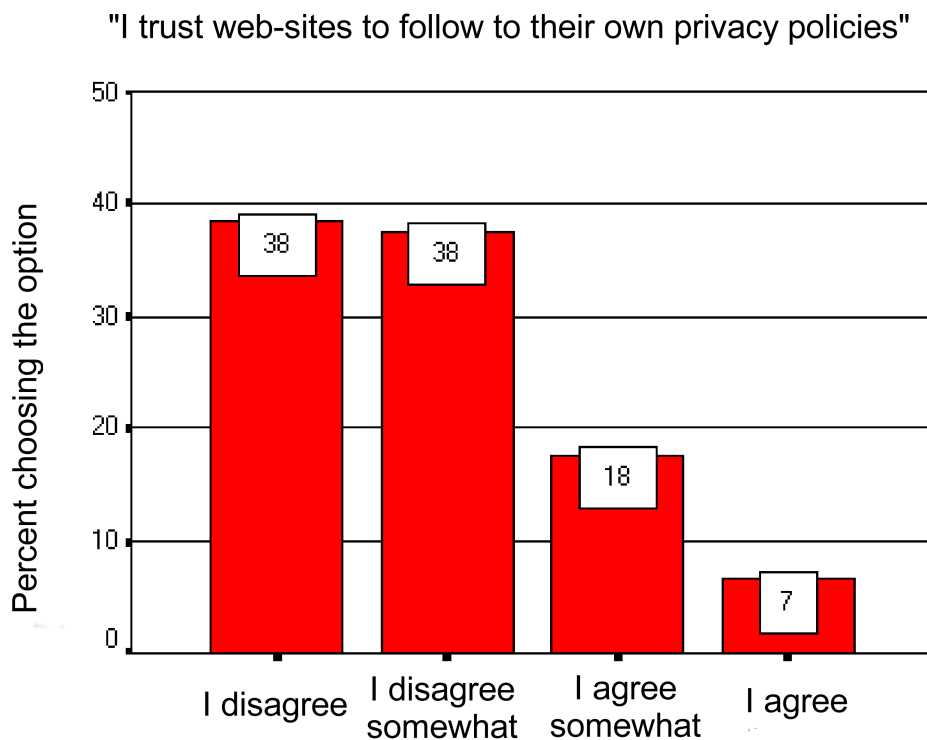
Figure 2 User Faith in web-site's compliance with their own privacy policies

Question
"I trust web-sites to follow their own privacy policies."

Key result
75% of the sampled users disagreed somewhat or fully. Only 25% (somewhat) trusted web-sites to stick to their own privacy policies.

This lack of trust is interesting because outside of existing regulation (such as in Germany), most web-sites are relatively free to determine the content of their privacy policies for themselves. Obviously, there is a lot of work left to be done in order to (re?)gain user trust on this matter.

5.3 Which are the preferred strategies for providing information in order to gain access to an incentive?

Users were asked to choose their preferred strategy for revealing information when they register in order to get an incentive. The three alternative strategies were determined in a pre-study. (Figure 3)

Question
"If you are only filling out an online-registration in order to receive an incentive, which strategy do you choose?

a) I report everything truthfully
b) I report some of it truthfully but falsify the information on questions that are too inquisitive.
c) I generally give false information.

Key result
The overwhelming majority of 65% will protect their privacy by providing false information in response to questions deemed too inquisitive. 25% say they report everything truthfully, 13% said they falsify everything reported.

The predominant strategy of falsifying part of the information as required is a serious problem. It makes tracking down low-quality information harder because those parts of a registration that can be checked against independent sources (i.e. Name and Address) may be accurate, whereas other parts of the registration will contain false information. The 13% who tend to give only false information are less of a problem, as they are more easily identified and can thus be removed.

"If you are filling out an online-registration in order to receive an incentive, which strategy do you follow?"
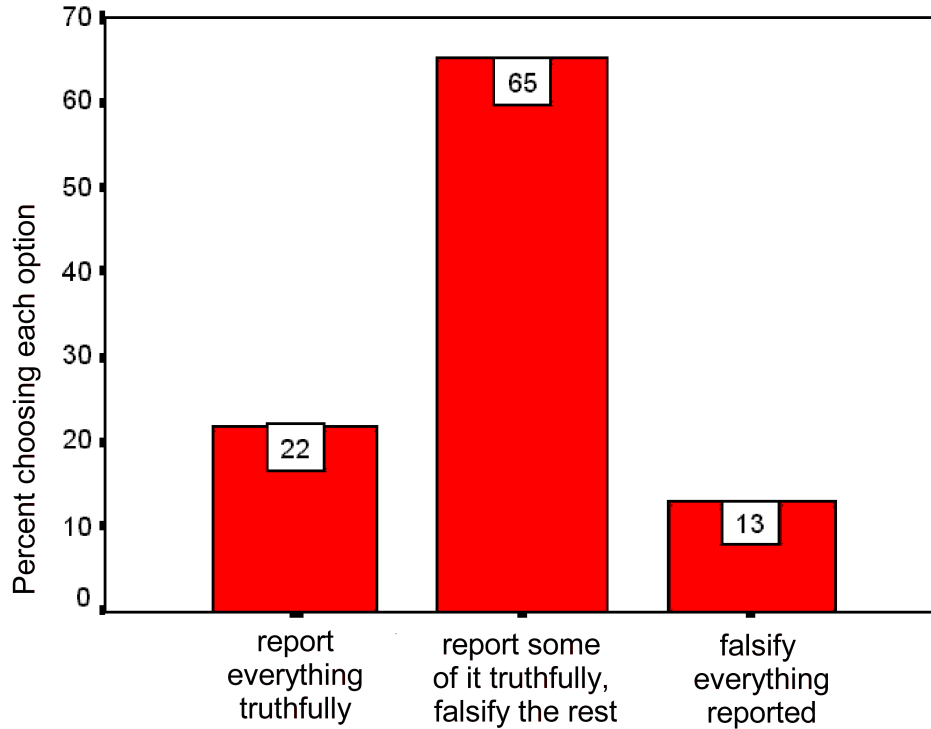
Figure 3 Preferred strategies for providing information as part of an online registration in order to gain access to an incentive.

5.4 What are the strategies for dealing with requests for information the user does not wish to provide?

Again we asked for user's strategy when dealing with trade-off decisions regarding online-registrations, this time focussing on the "points of truth": What happens when the user encounters a question that the site will force him to "answer" (in effect, except for an e-mail address, it can only force the user to enter something that *looks* sensible), but which the user does not want to answer truthfully. (Figure 4)

During the pre-study phase, three strategies were identified: a) *Quitting and forfeiting the offer*, b) *providing false information* and c) *first looking for an alternative source* for the incentive.
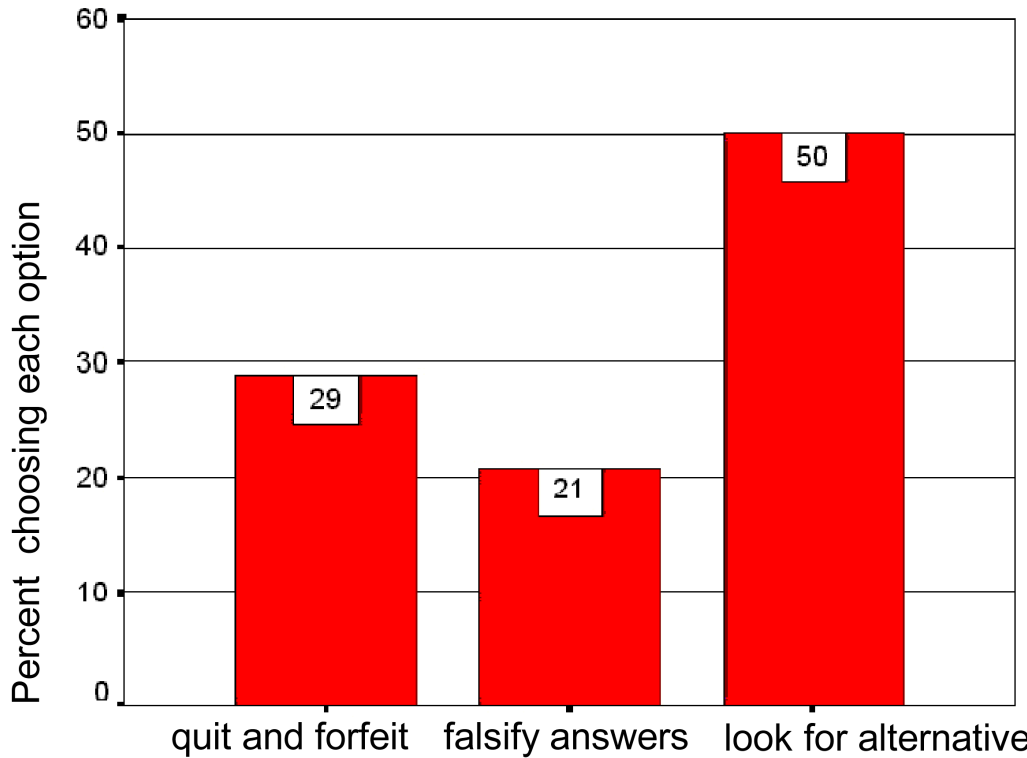
Question
"If, in the course of an online-registration for an offer you are interested in, you are asked for information you do not wish to reveal – how do you react?

a) I quit the registration process and forfeit the offer
b) I provide false information
c) I first look for an alternative source

"If, while registering for an offer, you are asked for information you do not wish to reveal - how do you react?"

Figure 4 Strategies for dealing with requests for information the user does not wish to provide

Key result

About 50% say they would first look for an alternative source for the offer, where the sensitive information is not required, 29% say they would forfeit the offer and 21% say they would provide false information on the spot. If this result is any indication of actual user behaviour, then web-sites need to take a much closer look – from a competitive point of view if nothing else – at what kind of information they *demand* and what kind they make *optional*. In a further question those who said they would first look at an alternative were asked what they would do if their search did not yield an alternative. Interestingly, the same proportion of "quit" vs. "fake information" was foud as in the first question. This suggests that the "quitting vs. faking" disposition and "deciding now vs. looking for an alternative" are two truly independent variables.
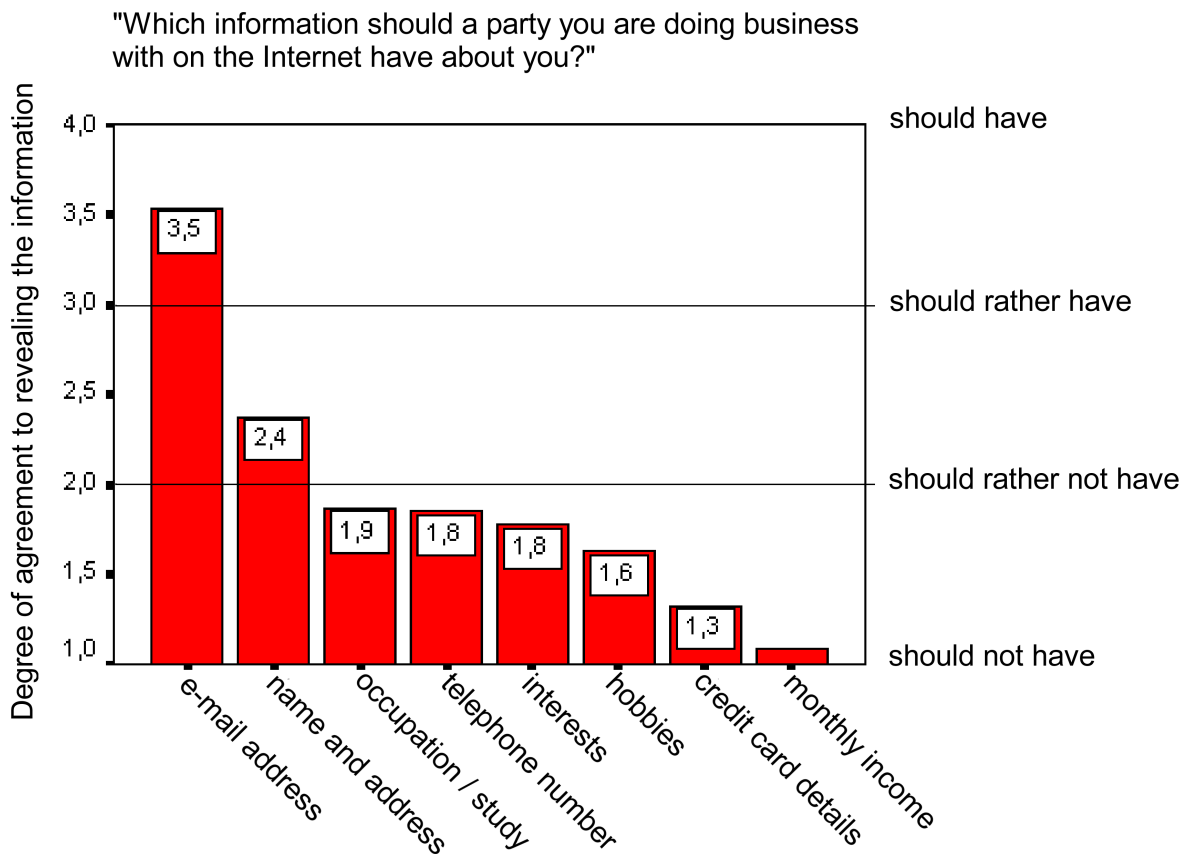
5.5 Which kinds of information are users prepared to reveal to someone they are doing business with online?

Finally, we asked which information users would willingly give to a business they are in contact with online, as the degree to which users voluntarily provide a certain kind of information should be be a good indication of the quality of that particular kind of information in online registrations.

Users were asked to rank for each information on a 4-item scale on whether they feel a business they are dealing with online should have that particluar kind of information about them. The ratings are: *4-should have, 3-should rather have, 2-should rather not have* or *1-should not have*.

Question
"Which kinds of information should a party you are doing business with on the Internet have about you?"

Figure 5 Which kinds information are users prepared to voluntarily reveal to someone they are doing business with online?

The kinds of information users were asked to rate were (ranked here from disclosure to non-disclosure preference, 4 meaning voluntary disclosure, 1 meaning non-disclosure):

1) e-mail-address   3.5
2) name and address   2.4
3) occupation / study   1.9
4) telephone number   1.8
5) interests   1.8
6) hobbies   1.6
7) credit card details   1.3
8) monthly income   1.1

Key result

These 8 kinds of information, which are also part of the Open Profiling Standard (OPS; Netscape 1997) and can therefore be assumed to be of some interest to marketers, have been ranked in order of the arithmetic mean of user's agreement to letting their business "partners" online have them. The most striking result is that the e-mail-address is the only piece of information that the surveyed users voluntarily offer. All the other pieces of information are already below the *point of indifference* of 2.5. While the revelation of "name and address" information is still somewhat tolerable (2.4), all the other kinds of information fall between "*should rather not have*" and "*should not have*". This striking preference demonstrates that web-sites need to be very clear about *why* they demand a particular piece of information from a user – and be able to communicate that need, for anything going beyond an e-mail-address and perhaps a traditional address is not voluntarily given to a party the user is *doing business with* online, never mind a totally unfamiliar company. In conjunction with the results in Figure 3, this can also be taken to imply that anything *other* than an e-mail-address (and perhaps a regular mail-address) is of highly dubious quality if it was collected in an online-registration's *compulsory* field and was not independently verified.

6 Outlook

With the growing popularity of the Internet, there is a shift of control over the communication away from companies and to their customers, as illustrated by rating-sites such as *epinions.com* and the popularity of the Cluetrain Manifesto (*cluetrain.com*). Against these changes, the importance of *privacy, trust* and *permission* in marketing communications will inevitably increase.

Current business trends such as *one-to-one* marketing and *personalization* of information, services and products on the internet require *accurate* information about customer interests and desires. While some of that information may be gathered implicitly by web-tracking or third-party information, much will depend on the accuracy of information explicitly gathered by means of an online-registration.

As this study has shown – and as proponents of permission marketing such as Godin (1999) demand – the issues of trust and privacy are the key problems to be adressed with respect to gathering high-quality information online. While incentives such as access to software,

information or sweepstakes can boost the *quantity* of registrations on a web-site, these incentives do not further the *quality* of the gathered information.

The results of this study suggest that maybe it is time for online marketers to take a step back and re-arrange their priorities to co-incide with prospective customers' needs. I.e., is it *really necessary* to know the name, street address, date of birth and day phone number that so many sites demand users tell them in exchange for a minimal incentive such as a free e-mail-address? Granted, the address information will allow a marketer to get a good guess at someone's average income, and may be used for tapping other sources of information. But examined closely, this craving for demographic information is the hallmark of traditional demographic segment marketing strategies that may be hopelessly outdated when compared with the options offered by this highly interactive medium, the Internet.

There are trade-off decisions that online marketers need to consciously make between the quantity of information demanded and the quality of information received. Would they rather have a lot of false information or a little accurate information? Is there a real need to have more than an e-mail-address to contact a prospective customer? Is there any realistic use in requiring a user to reveal his telephone number, other than a dreaded telemarketing call at dinner time? What value is in it for the user?

Privacy has been described as "The right to be left alone". Due to the highly interactive environment of the Internet, online marketers may be well advised to honor this wish by letting listening to what information their users are willing to share, maybe even under a pseudonym, rather than forcing them to fill out lengthy forms with information they would rather keep private. Enforcing compliance will result in users' taking measures to protect their privacy – at the expense of the quality of information in the marketer's database. With the user increasingly in control of the communication, companies should be actively working *with* prospective customers, not against them, if they want to prosper in the future. If we look at online reagistrations as a *value exchange* designed to initiating a long-lasting and mutually benificial relationship – rather than a quick means to generate addresses – out goal becomes that much clearer.

References

Fittkau, S. & Maaß, H. (1999). *W3B-Ergebnisband – W3B-Benutzer-Analyse Oktober/November 1998*. Hamburg: Fittkau & Maaß.

GameSpy Industries (1999). Mp3Spy Registration. [On-line] Available:
http://www.mp3spy.com/register/lost.shtml

Godin, S. (1999). Permission Marketing: Turning Strangers into Friends and Friends into Customers. New York: Simon & Schuster

Gräf , L.& Heidingsfelder, M. (1999). Bessere Datenqualität bei WWW-Umfragen – Erfahrungen aus einem Methodenexperiment mit dem Internet-Rogator. In: Batinic, B. et al. [Hrsg.]: *Online-Research: Methoden, Anwendungen und Ergebnisse*. Göttingen: Hogrefe.

GTRC – Georgia Tech Research Corporation (1998). GVU's 9th WWW User Survey: Web and Internet user Summary. [On-line] Available:
http:// www.gvu.gatech.edu/user_surveys/survey-1988-04/reports/1998-04-Use.html

Herzberg, F., Mausner, B., & Snyderman, B.B. (1959). *The Motivation to Work* (2nd ed.). New York: John Wiley & Sons.

Huff. D. (1954). *How to lie with statistics*. New York: Norton & Company

IMO – Information Market Obeservatory, European Commission (1995). ADMEDIA Project Report: The future of media and advertising. [On-line] Available:
http://www2.echo.lu/imo/en/final_ex.html

Netscape Communications Corporation (1997). The Open Profiling Standard (OPS). [On-line] Available: http://developer.netscape.com/ops/practices.html