

Using Online Reviews as Narratives to Evoke Designer's Empathy

Christiane Grünloh, Åke Walldius, Gerhard Hartmann, Jan Gulliksen

► **To cite this version:**

Christiane Grünloh, Åke Walldius, Gerhard Hartmann, Jan Gulliksen. Using Online Reviews as Narratives to Evoke Designer's Empathy. 15th Human-Computer Interaction (INTERACT), Sep 2015, Bamberg, Germany. Lecture Notes in Computer Science, LNCS-9296 (Part I), pp.298-315, 2015, Human-Computer Interaction – INTERACT 2015. <10.1007/978-3-319-22701-6_22>. <hal-01599656>

HAL Id: hal-01599656

<https://hal.inria.fr/hal-01599656>

Submitted on 2 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Using Online Reviews as Narratives to Evoke Designer's Empathy

Christiane Grünloh^{1,2}, Åke Walldius¹,
Gerhard Hartmann², and Jan Gulliksen¹

¹ KTH Royal Institute of Technology, Lindstedsvägen 3, 10044 Stockholm

² CUAS Cologne University of Applied Sciences, Steinmüllerallee 1, 51643 Gummersbach
{christiane.gruenloh, gerhard.hartmann}@fh-koeln.de
aakew@nada.kth.se, gulliksen@kth.se

Abstract. Gathering health-related data is quite easy, but visualizing them in a meaningful way remains challenging, especially when the application domain is very complex. Research suggests that empathy can facilitate the design process and that narratives can help to create an empathic encounter between designers and the prospective users. We conducted an exploratory quasi-experiment in order to explore whether narratives in form of online reviews are able to evoke designer's empathy when developing an online platform for a direct-to-consumer genetic testing service. The results suggest that the narratives can help designers to engage with and take the perspective of the prospective user, who is then represented in more detail. Lacking narratives from real people leaves the designers to their own imagination, which can lead to the use of rather abstract stereotypes that do not enable an understanding of the user, but affect the subsequent design decisions.

Keywords. Human-Computer Interaction, Empathy, Direct-to-Consumer, Genetic Testing, Health and Wellbeing

1 Introduction

Consumer products with regard to health and wellbeing are on the rise. The increasing availability and affordability of self-tracking devices and apps enable people easier access to their health-related data. Recently Topol (2015) compared the smartphone to the Gutenberg press, in that it might help to break boundary in medicine, because patients today are able to take a more active role (e.g. accessing various information sources or using health apps) [33]. Data can be collected by the hardware itself (e.g. data related to physical activities like daily steps, heart rate, distance, speed, duration) or manually tracked by the individual (e.g. nutrition, mood), and shared with others (e.g. PatientsLikeMe¹). Additionally, services like direct-to-consumer (DTC) genetic testing allow for easy access and exploration of personal genetic information. The boundaries between services with regard to health and wellbeing on the one hand and

¹ <http://www.patientslikeme.com/>, (accessed: January 17, 2015)

medical issues on the other can easily become blurred. The very same app could be used as a tool for managing health and wellness, but also for medical purposes. Although the data collection is quite easy, it has to be processed and visualized in a way that helps the individual to understand them in order to generate meaning.

According to the ISO 9241-210 user experience includes “all the users’ emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after use” [17]. Especially when it comes to services like genetic testing, the time span *after* the service has been used may become very important to consider. There are concerns that with “respect to asymptomatic individuals, [...] genetic testing may trigger an untoward psychological response, such as severe depression, anxiety, or even suicidal ideation” ([32], as cited in [15, p. 5]). Designing a system or service in this area can be very challenging, because it entails potentially complex data and addresses various customers with different motivations and backgrounds. Wright & McCarthy explored how empathy can facilitate the design process in terms of “knowing the user” [38]. Besides ethnographic approaches, which can be very time consuming, they considered also methods and techniques that involve empathic encounter without direct contact between designer and participant, e.g. the use of narratives.

In this paper we want to explore whether the reading of and the dealing with narratives in the form of online reviews is able to evoke or elevate the designer’s empathy with the prospective users of their design. Furthermore we are interested whether the reading of the afore-mentioned material helps designers to identify important experiences, which should be considered and addressed in the design of a system or service. We tried to answer these questions by conducting an exploratory quasi-experiment, where students were asked to develop an online platform for a direct-to-consumer genetic testing service. We regarded this case of application appropriate, because the service provides complex information, which cannot clearly be categorized as health, wellness, or medical data. Therefore it offers great potential for multi-dimensional discussions and application of various perspectives, precisely because the testing can entail both innocent and very serious results. In order to reduce the distance to people and enable empathy and understanding, we used narratives and personal accounts that have been published by real people on the Internet.

2 Background

2.1 Empathy and Related Works

The value of empathy and the opportunities it provides as a useful tool has been discussed on several occasions (e.g. in ergonomics [30], HCI [38], participatory design [19], product design [28], design research [22]). Dandavate et al. even go so far to say that the success of products will depend upon the degree to which researchers and practitioners learn how to empathize with the product users early in the development process [8]. Keen defines empathy as a “vicarious, spontaneous sharing of affect” which can be provoked “by witnessing another’s emotional state, by hearing about another’s condition, or even by reading” [18, p. 208]. This can be referred to as the *affective* approach, which considers one’s emotional response to the affective state of

the other [2]. Empathy is also often understood as the ability to ‘walk a mile in someone’s shoes’, i.e. to take the role of someone else or another’s perspective, which can be seen as the *cognitive* approach of empathy [2].

Researchers and designers make use of specific methods like for example Experience Prototyping or autoethnography, which support them to take the perspective of the user. *Experience Prototyping* aims to support designers, users, and clients to understand existing user experience and future conditions by engaging with the prototypes themselves [4]. Buchenau & Suri describe as an example project the “Patient Experience”, which addresses patients with chest-implanted automatic defibrillators. Due to the lack of first-hand experience by real patients, they wanted to recreate the essential elements of a personal experience, namely how it is like to be a defibrillating pacemaker patient. This was done by distributing pagers to the members of the design team as a proxy, where the pager represented defibrillating shocks at random times. The authors state that in this project empathy was promoted by the Experience Prototypes, but that this method should be seen as complementary to other design methods helping to understand other people's points of view [4, p. 432]. *Autoethnography* can be seen as related to Experience Prototyping. Here the researcher adopts the role of the participant as well in order to “understand and empathize with the experience mobile device user can face in difficult to access contexts” [25]. O’Kane et al. used this method to evaluate a wrist blood pressure monitor, because the non-routine situations they wanted to investigate were situations real users would be reluctant to disrupt for a study (e.g. holidays, festivities, etc.). They conclude that “for non-routine times it is an insightful method for challenging assumptions, gaining empathy with user experiences, and planning future user studies, including with mobile medical technologies” [25, p. 990].

Wright & McCarthy explore the relation between empathy and experience in Human-Computer Interaction (HCI) [38]. They see “empathic approaches as part of the broader pragmatist approach to experience”, because from “the pragmatist perspective, understanding an other or more specifically, ‘knowing the user’ in their lived and felt life involves understanding what it *feels like* to be that person, what their situation is like from their own perspective. In short, it involves empathy” (emphasis in original) [38, p. 638]. Questioning whether an empathic encounter without direct contact is possible, the authors explore narrative approaches that have been created by HCI practitioners, e.g. ethnographic vignette, character-driven scenarios. They conclude: “Using these methods in the spirit of enquiry and responsive understanding in which they were intended to be used may be sufficient to provide empathic understanding.” [38, p. 644]

Storyboards and narratives are a way to elicit empathy when direct contact to real users is not possible. McQuaid et al. used these as “customer surrogates” in order to understand their frustrating and pleasurable experiences with a national, public library and communicate those to stakeholders [23]. The storyboards were produced by user research specialists after they had acted as participants trying to fulfill a certain task. The authors consider these storyboards and narratives of select personas as effective techniques to help stakeholders to empathize with their customers. They believe that the stakeholders engaged more with the stories, because they were very realistic in the

sense that they included real people and pictures, and that the process was very concrete and visual represented in the stories [23].

Cooper introduced personas as “hypothetical archetypes” of actual users which are defined by their goals and “with significant rigor and precision“, although they are imaginary [7]. As concrete representations of target users, they help the design team to become more user focused by “putting a face on the user” and conveying information about the prospective users in ways that other artifacts cannot [27, p. 11]. These have been developed further by Pruitt & Grudin and used not only in the design team but for communication purposes to all project partners (i.e. developers, testers, writers, managers, etc.) [26]. Personas can help to make “assumptions and decision-making criteria equally explicit” and that without them certain decisions are made routinely “without recognizing or communicating their underlying assumptions” about how the product will be used by whom [26]. However, already the process of creating the personas helped them to make the assumptions about the target audience more explicit [26]. Personas make use of images and the act of finding the right image for the persona can already stimulate empathy with users [37].

With regard to narrative approaches it is emphasized that abstract representations of the individuals are counterproductive, and that it is important for the designer to “engage with the characters” and understand “their background, personality, intentions, and motives” in order to explore “how that person might respond to new situations and new technologies” [38, p. 642]. According to Nielsen, describing “the user as a rounded character” helps the design team to engage with the user with empathy, but the description must be based on “knowledge of actual users”; i.e. on facts and not fiction [24, p. 104]. Golsteijn & Wright use a portraiture approach, in which the story of each interviewed or observed person is told in a separate individual portrait, “within the context of use and staying true to the real user” [16, p. 301]. The authors believe that their approach “minimizes the risks of stereotyping and oversimplifying users and their experiences” and that these holistic descriptions of real users should be imported directly in the ideation process (unlike personas which are more generalizing / summarizing and created *before* ideation has begun) [16, p. 301f].

Researchers aiming for improvement of quality of health care services and patients’ experience developed the *experience-based co-design* (EBCD) approach, which also bases on narratives and personal stories [31]. This approach uses video narrative interviews with local patients and staff to identify opportunities for improvement from both perspectives. Due to the involved time and cost, this approach has been adapted to *accelerated experience-based co-design* (AEBCD), in which a national archive of patient narrative interviews² is used instead of interviewing and filming local patients [20]. Although national instead of local interviews were used, those films were considered “not perfect, but were ‘good enough’ to start the process of co-design” [20, p. 37f].

² The national archive is held by the Health Experiences Group in Oxford (HERG) and online available at <http://www.healthtalk.org/> (accessed: January 17, 2015)

2.2 Genetic Testing

Due to the decreasing cost of testing and sequencing of genomes, genetic testing is becoming more and more popular. For example the Department of Health launched a project to sequence 100,000 whole genomes from NHS patients by 2017 [13]. However, genetic tests are also sold direct-to-consumer (DTC), which has been criticized due to ethical, legal, and social issues [5]. Depending on the information these tests provide, it is not easy to draw a clear line in categorizing the test as a medical product or an information product. Companies like 23andMe for example offer at the same time tests for harmless traits (e.g. eye and hair color) and serious diseases (e.g. Alzheimer's Disease, Breast Cancer, etc.). Packaging together trivial and potentially life-changing tests can be problematic, especially when this serves as a means to get around regulation when only the fun part is advertised [35]. The U.S. Food and Drug Administration (FDA) considers the service by 23andMe as "intended for use in the diagnosis of disease or other conditions or in the cure, mitigation, treatment, or prevention of disease, or is intended to affect the structure or function of the body" [34]. While the company is still waiting for an approval in the U.S., the service was approved in the UK, because the UK Medicines and Healthcare Products Regulatory Agency (MHRA) considered the service as an information product instead of a medical product [14]. A detailed discussion of the ethics involved in DTC genetic testing is beyond the scope of this article. For a more comprehensive discussion on the benefits and concerns regarding DTC genetic testing see for example [29] and [5].

2.3 Doctor-Patient Relationship

Due to technology and other developments in healthcare the relationship between patients and the healthcare professional is changing. The more traditional relationship, also known as the "paternalistic model", assumes that the doctor can judge patients preferences, that both have the same goals, and that only the doctor has the expertise to determine how to proceed [9, p. 171f]. Moreover, this model also assumes that "the doctor will make the best treatment decision for the patient and can do so without eliciting personal information from the patient or involving him or her in the decision making process" [6, p. 781]. This has changed in that sense, that today the autonomy of patients is supposed to be respected and some alternative models emerged, e.g. the *engineering model* (patient is the sole decision maker, physician only gives advice), the *collegial model* (recognizes the imbalance of knowledge and views between patients and providers; sees them as equal partners), the *contractual model* (shared decision making with contributions by both patient and physician) [36, as cited in 9, p. 172]. Emanuel & Emanuel outline four models of physician-patient interaction (paternalistic, informative, interpretive, and deliberative model) and compare them with regard to patient values, physician's obligation, conception of patient's autonomy, and conception of physician's role [10]. The authors recognize that different models may be appropriate under different clinical circumstances, but emphasize that the paternalistic model is justified in rather limited circumstances (i.e. emergencies), because "it is no longer tenable to assume that the physician and patient espouse similar values

and views of what constitutes a benefit” [10, p. 2224]. Nowadays interventions for patient empowerment have been carried out, which aim to “increase the patient’s capacity to think critically and make autonomous, informed decisions” [1, p. 279]. This is also supported by governments in Europe, which are promoting the “expansion of eHealth over the past years, arguing that this development enhances patient participation, empowerment and cost efficiency” [11, p. 1].

3 Method

In order to investigate whether the reading of narratives and reviews is able to evoke or elevate empathy for people, to the extent that it may change the design of a system or service, we conducted an exploratory quasi-experiment with a between-subject design. Two groups of students were asked to develop an online platform for publishing genetic test results for a direct-to-consumer genetic testing service. Since the students were not familiar with these services, we could assume that they were not aware of the controversies involved. The experiment considered as independent variable the material the groups were provided with, which for the experimental group additionally contained personal reviews written by actual customers of a direct-to-consumer genetic testing service. We regarded the dispositional empathy (i.e. empathy as a character trait) as a possible confounding factor. If for example the participants assigned to the control group are already very empathic and therefore don’t necessarily need to read narratives as a means to empathize with a person. To control the confounding factor, we conducted a pretest in which the participants had to complete the Empathy Quotient (EQ), a self-assessment instrument developed by [2]. The results of the EQ served as a means to assign the participants to the two groups to achieve comparability.

3.1 Pilot study

The initial experiment design was tested with two students, who didn’t participate in the main study, in order to evaluate the feasibility of the study design and whether the material is comprehensible for the participants. It transpired that the task description and instructions had to be clearer and partly reworded. The students struggled when reading the English narratives. Therefore we decided to translate the customer reviews into German to make sure that they are fully understood.

3.2 Participants

The participants were recruited via student mailing lists and announcements on the University’s wiki. The experiment comprised 14 advanced students of Media Informatics (8 male, 6 female; 10 Bachelor students, 4 Master students; age: 22-35 years). None of the participants have carried out a genetic test in the past.

3.3 Procedure

Due to the small sample (N=14) random sampling was not appropriate. To reduce the effect of the confounding factor *dispositional empathy*, the groups were divided by help of matching. All participants completed the EQ questionnaire a few days prior to the study. We then assigned the individual participants to achieve comparable groups regarding the EQ (see table 1). Both groups equally consisted of 7 participants and we also managed to have equal number of female/male, Bachelor/Master students in each group.

Table 1. Median, Means and Total EQ

Group	Median	Mean	Total
Control Group	41	40.857	286
Experimental Group	42	40.857	286

On the day of the study each participant received a consent form (and a personal copy), in which the process of the study was explained, as well as risks, advantages, how the study was recorded, how we deal with the data, and of course their right to withdraw from the study at any time. After the participants signed the informed consent, the study started with an introduction to the task and the procedure in general. We introduced the task description verbally and informed the participants, that we would separate them in two different groups, in which they then could work on the task. We then showed them four small video clips in order to make the participants familiar with the nature of genes, SNPs (Single Nucleotide Polymorphism), and phenotypes. The videoclips have been produced by 23andMe, published on youtube.com³ and explain the topics in a rather playful and almost trivializing way. Taken the background music and the style (an animated cartoon) into account, these clips seem to target rather children than adults. However, it was important for us to adhere to the marketing strategy of those companies.

After showing the video clips, we separated the two groups and assigned them to two different rooms. Each group received a notepad for the group document and each participant also received a notepad for his/her individual notes. In addition every participant received a printed task description including the instructions, the procedure, and an excerpt of the diseases and conditions the service includes in its genetic analysis. This information was also taken from the 23andMe website⁴, but shortened and included the categories: Ancestry Composition, Disease Risk, Drug Response, Traits, and Carrier Status. The material of the experimental group was supplemented by five personal reviews. In order to provide the participants in the experimental group with rich narratives that might enable the identification with the user of the service, we

³ The videos are available at http://youtu.be/ubq4eu_TDFc?list=PLF9969C74FAAD2BF9 (last access: January 17, 2015). We showed the clips 1 to 4, but didn't view the first seconds where the company name is displayed.

⁴ See <https://web.archive.org/web/20130424163612/https://www.23andme.com/health/all/> (last access: January 23, 2015)

extracted real customer reviews from amazon.com⁵. Only little information was deleted or altered (e.g. we substituted the name of the company with a place holder). The reviews were chosen because they gave some hints regarding the motivation of using the service (e.g. one person was adopted and wanted to learn about diseases that run in the family or another wanted to know about the chances his kids develop schizophrenia). There were some negative comments within the reviews, but also some positive ones. As mentioned earlier, we translated the reviews into German to increase the comprehensibility. The experimental group got instructed that they would have to read all of the material, but they are not bound to use it in their design. We wanted to decrease the possibility that the participants regard the purpose of the material as something that they *have to* analyze in depth. After reading the instructions both groups were observed but not guided through their design process. Only questions regarding the procedure (e.g. how much time they have left) were answered by the observers. After approximately one and a half hours, the students were asked to stop their activities. After a short break, both groups presented their results to all participants. Finally we debriefed the participants and answered pending questions (e.g. regarding the purpose of the questionnaire they completed in advance).

3.4 Data Collection and Analysis

The data comprised video recordings of the group work (about 2 hours per group) and the presentations, the group document written by the individual group secretary, individual notes of the participants, pictures from the group's output (e.g. from brainstorming, clustering) and observants' notes. Each group was recorded using the in-build camera and microphone of a MacBook Pro. The data was transcribed verbatim afterwards using MAXQDA11 for Mac⁶. The transcription was then analyzed using inductive Thematic Analysis as described in [3] to identify recurrent themes in the data; within and between the groups. After familiarizing with the data, 40 initial codes were generated by systematically working through the entire data set, one group at a time. The initial codes identified features like emotions, process (use of analogies, scope of the task), concerns (consequences, risks, comprehensibility for laypeople, legal issues, misuse, incorrect results, serious results), design (e.g. customer service, support, usability, functionalities, change of testing process, visualization of results), perspectives addressed (e.g. health insurance, physician, genetic testing company, design company, users, motivations), values (e.g. ethics and morals, freedom of choice, paternalism, data privacy, security). Afterwards the entire data set was reviewed with regard to the codes to examine coherence of the codes between the groups. Then the extracts were reviewed on code level, i.e. all extracts of both groups corresponding to a specific code were reviewed in order to analyze the codes. Due to the exhaustive set of data extracts and codes, extracts were exported from the software tool MAXQDA for further investigation and inspection, to collate the coded

⁵ See <http://amzn.com/B002QPR852> (last access: January 23, 2015)

⁶ For further information see <http://www.maxqda.com/products/maxqda>, (accessed: January 17, 2015)

extracts within a code, and to “consider how different codes may combine to form an overarching theme” [3, p. 89]. In accordance to [3] this phase re-focuses the analysis at the broader level of themes rather than codes. Visual representations like mind maps and tables were very helpful to sort the codes into themes and to identify similarities and differences between the groups. By means of reading, categorizing and reviewing the codes and extracts for each theme repeatedly, the themes could be developed iteratively and will be described in detail in the next section.

4 Results

Our analysis showed that I) the groups share deep concerns regarding the DTC genetic testing service; II) the groups differ in how they represent the prospective users; III) they differ in how they deal with their concerns with regard to their representation of users; IV) the design decisions differ accordingly. These themes will be elaborated in detail in the following sections. For references to the data and quotations, the group is indicated as EG (experimental group) or CG (control group) and the number of the extract is added. The excerpts presented from the transcripts have been translated from German to English by the first author.

4.1 Shared Concerns

Both the experimental group (EG) and the control group (CG) shared several concerns with regard to the introduced DTC genetic testing service:

Revelation of serious results. Both groups identified early on that “for some people a genetic test is a very serious matter; for others it’s more a gimmick” (EG:109). While the groups considered the rather harmless information (like Ancestry Composition, Drug Response, and Traits) to be revealed straightaway, information regarding Disease Risk and Carrier Status should be taken seriously and processed differently.

Both groups struggled with how to reveal such sensitive information and if the Genetic testing company is generally allowed to share such information with people. Both share their concern that the results might cause panic, psychic stress, or upset a person, when revealed via an online service. Both groups indicate on several occasions that they want to include some sort of psychological support, e.g. psychological counseling, reference to support groups, support by phone, or aftercare in general.

The groups share the concern with regard to undesired consequences, but the type of consequence differs between the groups. CG referred to rather extreme consequences due to how people would *react* after learning about the results (e.g. that a person would commit suicide, give up their child for adoption when the results are bad, or even perform an abortion). EG considered consequences in terms of what the results might *mean* for the person in their future (e.g. an unwanted result in a paternity test, or having a high probability regarding Alzheimer’s or cancer).

Comprehensibility. Both groups identified that the information is quite complex and that laypeople probably need additional information and a comprehensible visualization. Both also realized that they as designers also struggle to understand the specific terms, which they feel is necessary in order to structure and cluster the data for laypeople. The experimental group went one step further in discussing that not only the lack of knowledge, but also the existence of prior knowledge (i.e. experts) should be addressed, because those “would probably rather understand professional jargon” (EG:314).

Data access by third parties. The groups identified potential interests in the data by third parties, e.g. insurance companies, physicians, bone marrow database, or research (CG); employer, anyone else besides the donor (EG). In this regard, both groups also discussed the risk that an unauthorized person might send a sample of someone else.

4.2 Representation and attitude towards users

The way the two groups talked about the users differed noticeably. The representation of users remained very abstract in the control group. When they discussed who would want to do such a genetic test, they referred to “people who panic to get sick” (CG:43); “curious” (CG:44, 179), “who have been adopted” (CG:45, 48), “who consult online docs” (CG:49), “overcautious” (CG:172), “doctors could use it for their patients” (CG:173), “hypochondriac” (CG:307,370), “parents who want to test the DNA of their children” (CG:503). The representations remained rather stereotypic without deep discussion regarding the characteristics of the person or further motivational aspects.

The experimental group didn’t represent the users in such a stereotypic way. The discussions evolved here rather around the underlying motivations, interests, characteristics, and needs of people, e.g. “What would the customer, who doesn’t know exactly why he actually makes the test... what could be important to him and what would he want to know” (EG:111). The discussions with regard to a person’s background and characteristics were in all more detailed: “adopted child, who wants to know about their medical history and origin of the family” (EG:49); “in case a disease runs in the family, and one wants to know if one is affected” (EG:64), and with reference to a review: “Right here with schizophrenia. That he knows he does not have it so he’s not schizophrenic. But he’d just like to know whether his children could get it.” (EG:260); furthermore, a person might be interested only in certain information like origin of family - not more (EG:339); “people who have a serious disease” (EG:407); “users with shortcomings” which the student’s would like to address (EG:313); people without specific domain knowledge (“laypeople”), and who might need some anecdotes (EG:314); and “experienced persons” who require professional jargon (EG:314).

Besides the identification of motivations in the reviews, the students of the EG also wanted to learn about the person’s motivations, special needs, and disabilities (e.g. blindness) through some kind of pre-test, which would then lead to an individualized presentation (EG:308). Later in the process they also stepped back again: “We already started with the requirements, even though we didn’t really address the users. Who are

those people?” (EG:401), and further: what are their motivations (EG:403) and what additional information might be important to inquire up-front (EG:409) that could be used for individual representation (EG:414).

4.3 Paternalism vs. Autonomy

As already mentioned, the students struggled when it came to the revelation of serious results. The control group took quite early a rather paternalistic approach, in which the access to some results is denied. “I just had an idea... there are these online tests, so that you might ask some basic questions in advance ... and then you say: No, sorry. You are not getting the results. Get in touch with our doctors or whatever. Something like that.” (CG:26). Initial questions arose regarding the legal situation and if it is allowed to tell people their results “just like that” (CG:58), or whether some aspects would be excluded, a question which was met with the counter argument “How so? You always have a right to information” (CG:61). However, this wasn’t considered further after a student used an analogy: “Yes, but for example I know with Parkinson tests, they have to go through [psychologic counseling] and so forth before they learn, if they have it or not. And if the psychologists detect that the person is too unstable to learn that, then they won’t. That’s why I can’t imagine that all this information can be released just like that.” (CG:62) The group then developed their idea further that the service would offer two packages, where the one with Disease Risk and Carrier Status would be only available in cooperation with a doctor. They planned to grant only the specific doctor access to this kind of data (CG:76).

Some of the students in the control group noticed, that this would change the company philosophy, because it is a *direct-to-consumer* service after all (CG:70,83) and that changing the process the process might be bad for their business (CG:110, 112). They raised concerns whether they were allowed to make these kind of decisions (CG:151). However, in the end they considered it as important and in the person’s interest that the doctor is involved (CG:286): “the most important things that we really have in mind... what we are talking about the whole time, is to protect the user from himself and from the information he could get.” (CG:264). Moreover, not only the design team is able to decide for the person, the doctor can do that as well: “Then the doctor can, ... perhaps he knows his patient well... he can perhaps say directly: No, no, no, you don’t want to know the disease risks” (CG:280).

When one student asked, “If this is the decision of the customer? [...] who wants to get the results” (CG:288, 292), he was quickly overruled: “Well, bad luck, then go somewhere else” (CG:289); “Yes, but it’s for his own safety.” (CG:293). One student elaborated further on the protection with an analogy: “There are many things that are made for the protection of all of us, for example that we must fasten our seatbelt, etc. If I fast my seatbelt or not is still my own decision, but basically it is said you need to buckle up while driving, otherwise you could fly through the windshield.” (CG:296) The argument that the person could sign an informed consent was also overruled, because “But then he can also be quite unstable or so... I think that one always believes: I can take it, I can manage it.” (CG:300). The control group decided in the end that every person has to register and if they are interested in Disease Risk or Carrier Status, they have to cooperate with their doctor, who also has to sign a form.

The experimental group also had the idea to involve doctors or experts who might initiate psychological counseling (EG:115), identify the person (EG:117), or interpret certain test results: “bad news, yes... these diseases etc. that they are best interpreted by physicians” (EG:120). One group member early raised the concern that this would “depart from the business idea that they have” (EG:118). The group then changed the involvement of experts to be on a voluntary basis (EG:126) and that they would make recommendations who to contact (EG:131). The question if it is allowed or if someone is entitled to reveal such serious information in an automatized way (EG:132) and if this is ethical were met with “Yes, but he wants to know” (EG:133) and “He pays for it. He wants that.” (EG:135). Unlike the control group they dismissed ideas with paternalistic tendencies for the sake of respecting the person’s decision and gave the individual person and their perspective priority.

4.4 Design decisions

Changes with regard to the testing process. Although both groups considered to change aspects in the testing process of the DTC service, the underlying motivation differed and can be related to the previous theme. Both groups consider to include some kind of pre-test or application process. This was motivated in the CG in order to determine whether the person can handle this information. The pre-test was later obsolete, because according to their final design deliberations access to serious information was only granted through the doctor, who would then determine whether the person can handle it or not. The pre-test was considered in EG in order to determine an interest in specific information due to previous conditions (e.g. specific drug responses due to a certain disease) and to identify special needs of the person, which would then be addressed in the design; e.g. if the person is visually impaired.

Both groups considered to involve a doctor or an expert and mentioned that this could ensure the authenticity of the sample sender. However, this was not the initial motivation for the involvement, which differed between the groups. While the CG wanted to involve the doctor in order to reveal serious information in general, the EG wanted the doctor or expert to give *additional* information with regard to the results, interpret serious results, and initiate psychological counseling if necessary. Additionally, in the end only the control group kept the idea in their final design deliberations.

Usability. In line with the rather paternalistic approach, the CG tended to overload people with information in order to prevent them to sign up for this test thoughtlessly: “Rather put too much information than having the user click through it too easily” (CG:206); “I’d rather see to it that... that with the registration process, that in a sense scruple is generated” (CG:328). They recognized that it requires extra effort to undergo for example some sort of personal identification procedure (CG:329). They also considered that expressing additional consent that the person is aware of legal issues by an extra click might be contrary to usability, but they wanted to include that anyway “Just to make sure” (CG:438). The EG didn’t mention usability aspects explicitly, but discussed that they want to provide individualized visualizations with regard to the specific needs (e.g. auditive, textual, graphical, or a personal contact) and accord-

ing to the person’s specific interests: “...not some kind of heap of graphics and data, but to say, well, this is something important or this may be of interest for you...” (EG:111).

Functionalities. The CG elaborated several design dimensions with regard to safety aspects. This addressed the system’s security (restricted access, authentication, risk of having data online, ensure protection on a personal and data level). With regard to protection, the group wanted to make sure, that doctors will be involved, when health related results are to be revealed. Their solutions included an identification process and some kind of key encryption procedure to protect the data. The functionalities discussed in the CG dealt with Login, identification procedure, encryption, PDF export, anonymization, and verification. The EG focused not that much on technical solutions, although they also discussed some kind of personal identification procedure. Their design addressed functionalities like contact options for bi-directional communication with the company, the opportunity for ratings and reviews, provide communication platform between persons, filtering the results based on a person’s interest and data.

Visualization. The main task for both groups was to develop a platform that should be used to publish the genetic testing results. The outcome with regard to the visualization differed noticeably. The CG excluded the two categories Disease Risk and Carrier Status completely and didn’t discuss the visualization of the remaining categories in detail. They wanted to show certain information in a table on demand and if desired a person can get further information on a specific topic through hyperlinks. The EG discussed extensively how they want to visualize the information: e.g. visual and auditive, in a sensitive way, depending on the individual category (in a cheerful or serious manner), using some kind of color scheme (including the consideration of different cultural connotations), using different modalities (e.g. tree, table, list), using metaphors (e.g. traffic light, body parts, globe, maps).

4.5 Group presentation

Due to the time constraints, the ill-defined character of the design task, and the consequential extensive discussions within both groups, the final presentation of their design focussed on specific aspects they would address when developing the platform. Table 2 gives an overview about the aspects the groups presented:

Table 2. Design aspects presented by the groups

	Control Group (CG)	Experimental Group (EG)
Access to results	only access to <ul style="list-style-type: none"> • Ancestry Composition • Drug Response • Traits 	access to all categories

Doctor Involvement	mandatory for accessing <ul style="list-style-type: none"> • Disease Risk • Carrier status 	optional, but recommended for serious results
Visualization	<ul style="list-style-type: none"> • Visualization of results regarding Ancestry Composition, Drug Response, Traits • clear textual visualization • hyperlinks with further information 	<ul style="list-style-type: none"> • visual and/or auditive • sensitive way depending on category • harmless results in a cheerful way • address different user groups with special needs • use of metaphors (maps for ancestry; use of colors; body parts; traffic light metaphor for drug response) • comprehensible for laypeople and experts • prevent misunderstanding • use table, lists, hyperlinks
Other Requirements / Aspects	<ul style="list-style-type: none"> • Security • Data protection • Access restriction • Registration, unique ID, personal identification, • Export of data regarding disease risk and carrier status to doctor, who will discuss them with user • provide extensive information up front about what the service offers in order to create confidence • give information regarding counseling centers • option to share data with research, bone marrow databases • information regarding legal issues 	<ul style="list-style-type: none"> • High priority for motivation of users • Customer service • ensure transparency to create confidence • Trust • Data protection • User identification • prevent data misuse • dealing with consequences, offer a service to connect users to psychological counselor • enable import of data from insurance company / doctor (e.g. allergies to drugs)

5 Discussion

Both groups started with the same concerns regarding the type of service they were asked to develop an online platform for. How the groups discussed and talked about the people they were designing for differed in that the CG remained abstract and used rather stereotypical description of the persons. And as Nielsen states: “As the stereotypes will function as a mental picture they will never enable an understanding of the user.” [24, p. 104] The EG on the other hand described the users in more detail and took into consideration their background, interests, motivation, and needs. Although we can’t ‘measure’ quantitatively whether empathy has been evoked or elevated by the narratives, this might be a strong indicator. Like Nielsen stated: “To describe the user as a rounded character brings a focus on the user into the design process. It helps the design team to engage with the user with empathy, thereby remembering the user all the way through and remembering that the design is for a user.” [24]. The mental representation of the prospective users in the experimental group is likely to be influenced by the provided material (i.e. the narratives), because the EG referred to specific aspects of the narratives on several occasions.

The lack of empathy and/or the use of a rather stereotypical representation might have led the CG to follow a paternalistic approach, which then influenced the following discussions and design decisions. The stereotypical representation of the users can to some extent even be seen as preexisting individual bias (e.g. the “suicidal”, the “unstable”, the “hypochondriac”), which “has its roots in social institutions, practices, and attitudes” and “can enter a system either through the explicit and conscious efforts of individuals or institutions, or implicitly and unconsciously, even in spite of the best of intentions” [12, p. 333f]. The long tradition of the rather passive patient might have led to an unconscious bias in that they have to be protected and (in line with the paternalistic model) that it is appropriate to “spare patients the worry of decision making” [9, p. 172]. It is not our intention to imply that these systems should be build regardless of any concerns and that the user would themselves be responsible for any consequences. Rather to the contrary, we agree with Löwgren & Stolterman, that the “responsibility for what is created is fully in the hands of the creator - the designer“ [21, p. 4]. However, we think that making decision for the prospective users (e.g. “protecting him from himself”) without their involvement and hereby taking a paternalistic approach conflicts with recent attempts in terms of patient empowerment to increase the autonomy of people with regard to health and wellbeing.

We also noticed that the control group used analogies more often than the experimental group and lost themselves in discussions more often, whereas the EG managed to get back on topic faster. We cannot eliminate the influence of group dynamics, but one explanation could be that the EG had some material at hand, so that they did not need that much imagination and storytelling to get engaged in the task. This is in accordance to the work by Golsteijn & Wright, in which “the portraits acted to provide depth and focus, because rather than thinking about ‘anything’ we could think about the needs of one specific person – a real person – and what could be designed for this person” [16, p. 312].

6 Conclusions

In this paper we wanted to explore whether the use of narratives by real persons evokes or elevates the designer's empathy with the prospective users. Due to the sample size and the very design of the study (i.e. its explorative character, a design task as object of study within a restricted timeframe, and students as participants), the generalizability of the results is limited. However, based on the thematic analysis we conclude that providing additional material containing narratives from real people can help the design team to engage with and take the perspective of the prospective user, who is then represented during the discussions and in the design in more detail. Lacking narratives from real people leaves the designers to their own imagination, which can lead to the use of rather abstract stereotypes that affect the subsequent design decisions. However, it would be interesting to investigate, whether a study with professional designers with several years of experience would show similar results.

Although direct contact with real people is preferable for enabling an empathic encounter and to be able to "walk a mile in their shoes", this might not always be possible, e.g. due to time and budget constraints. However, today many people share their personal stories with others online; stories written by people in their very own words. These stories are therefore based on facts not fiction and might help designers to engage with the prospective users. Similar to other methods (e.g. experience prototyping) using online reviews as narratives should be seen as a complementary activity, for example before and during the ideation process, to avoid an early oversimplification of users. Based on our results we believe that these stories help to look beyond the initial preconception and to get an idea what it might be like to be the other. Personal narratives can help to get rid of stereotypes, preexisting biases, and fixations in our heads, in order to make room for people's voices. Therefore, it would be interesting to explore further, whether this could even be enhanced when designers are giving the task to create personas based on these stories, because the creation process as such has been found to make assumptions about the target group more explicit. Furthermore it would also be interesting to not only make use of stories in a textual form, but for example published videos of real people. In the case of Genetic Testing videos from people talking about their DTC genetic testing results⁷ or videos from the aforementioned archive of patient narrative interviews with regard to their experience with genetic testing⁸ could be applicable.

Acknowledgements. We would like to thank all participants of the study for their time and commitment. Furthermore, we thank the anonymous reviewers for their detailed comments and thoughtful suggestions on our submission, which were very considerate and helpful.

⁷ for example <https://www.youtube.com/watch?v=QqQpYEA9mpM> (accessed: January 17, 2015)

⁸ for example <http://www.healthtalk.org/peoples-experiences/nerves-brain/carers-people-dementia/genetic-testing> (accessed: January 17, 2015)

References

1. Anderson, R.M., Funnell, M.M.: Patient empowerment: Myths and misconceptions. *Patient Education and Counseling*. 79(3), pp. 277–282 (2010).
2. Baron-Cohen, S., Wheelwright, S.: The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders* 34(2), pp. 163-175 (2004)
3. Braun, V., Clarke, V.: Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3(2), 77-101 (2006)
4. Buchenau, M., Suri, J.F.: Experience prototyping. In: Proceedings of the 3rd conference on Designing Interactive Systems: processes, practices, methods, and techniques, pp. 424-433. ACM, New York (2000).
5. Caulfield, T., McGuire, A. L.: Direct-to-Consumer Genetic Testing: Perceptions, Problems, and Policy Responses. *Annu. Rev. Med.* 63(1), pp. 23-33 (2012)
6. Charles, C., Whelan, T., Gafni, A.: What do we mean by partnership in making decisions about treatment? *BMJ : British Medical Journal*, 319(7212), pp. 780–782 (1999)
7. Cooper, A.: *The Inmates are Running the Asylum*. Macmillan (1999)
8. Dandavate, U., Sanders, E. B. N., Stuart, S.: Emotions Matter: User Empathy in the Product Development Process. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting 40(7), pp. 415–418. Sage Publications (1996)
9. Deber, R.B: Physicians in health care management: 7. The patient-physician partnership: changing roles and the desire for information. *CMAJ: Canadian Medical Association Journal*. 151(2) (1994)
10. Emanuel, E.J., Emanuel, L.L.: Four models of the physician-patient relationship. *JAMA*. 267(16), pp. 2221–2226 (1992)
11. Erlingsdottir, G., Lindholm, C., Ålander, T.: eHealth services, patient empowerment and professional accountability-An empirical study on the changing patient-doctor relationship in the digital world. In: International EIASM Public Sector Conference, pp. 1-21 (2014)
12. Friedman, B., Nissenbaum, H.: Bias in computer systems. *ACM Transactions on Information Systems (TOIS)*, 14(3), pp. 330-347 (1996)
13. Genomics England <http://www.genomicsengland.co.uk/the-100000-genomes-project/> (accessed: January 17, 2015)
14. Gibbs, S.: DNA-screening test 23andMe launches in UK after US ban. *The Guardian* (2014) <http://www.theguardian.com/technology/2014/dec/02/google-genetic-testing-23andme-uk-launch> (accessed: January 17, 2015)
15. Goldman, J. S., Hahn, S. E., Catania, J. W., LaRusse-Eckert, S., Butson, M. B., Rumbaugh, M., Strecker, M. N., Roberts, J. S., Burke, W., Mayeux, R., Bird, T.: Genetic counseling and testing for Alzheimer disease: Joint practice guidelines of the American College of Medical Genetics and the National Society of Genetic Counselors. *Genetics in medicine* 13(6), 597-605 (2011)
16. Golsteijn, C., Wright, S.: Using Narrative Research and Portraiture to Inform Design Research. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013, Part III. LNCS, vol. 8119, pp. 298-315. Springer, Heidelberg (2013)
17. ISO 9241-210: Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems (2010)
18. Keen, S.: A Theory of Narrative Empathy. *Narrative* 14(3), pp. 207–236 (2006)
19. Lindsay, S., Brittain, K., Jackson, D., Ladha, C., Ladha, K., Olivier, P.: Empathy, Participatory Design and People with Dementia. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 521-530. ACM, New York (2012)

20. Locock, L., Robert, G., Boaz, A., Vougioukalou, S., Shuldham, C., Fielden, J., et al.: Testing accelerated experience-based co-design: a qualitative study of using a national archive of patient experience narrative interviews to promote rapid patient-centred service improvement. *Health Services and Delivery Research* 2(4) (2014)
21. Löwgren, J., Stolterman, E.: *Thoughtful Interaction Design: A Design Perspective on Information Technology*. MIT Press (2004)
22. Mattelmäki, T., Vaajakallio, K., Koskinen, I.: What Happened to Empathic Design?. *Design issues*, 30(1), pp. 67-77. MIT Press (2014)
23. McQuaid, H.L., Goel, A., McManus, M.: When you can't talk to customers: using storyboards and narratives to elicit empathy for users. In: *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces*, pp. 120–125. ACM, New York (2003)
24. Nielsen, L.: From user to character: an investigation into user-descriptions in scenarios. In: *Proceedings of the 4th conference on Designing Interactive Systems: processes, practices, methods, and techniques*, pp. 99-104. ACM, New York (2002)
25. O'Kane, A. A. Rogers, Y., Blandford, A. E.: Gaining empathy for non-routine mobile device use through autoethnography. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 987-990. ACM, New York (2014)
26. Pruitt, J., Grudin, J.: Personas: practice and theory. In: *Proceedings of the 2003 Conference on Designing for User Experiences (DUX '03)*. ACM, New York (2003)
27. Pruitt, J., & Adlin, T.: *The persona lifecycle: keeping people in mind throughout product design*. Morgan Kaufmann (2006)
28. Segal, L. D., Suri, J. F.: The empathic practitioner: Measurement and interpretation of user experience. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 41(1), pp. 451-454. Sage Publications (1997)
29. Su, P.: Direct-to-Consumer Genetic Testing: A Comprehensive View. *The Yale Journal of Biology and Medicine* 86(3), pp. 359–365 (2013)
30. Suri, J. F.: The next 50 years: future challenges and opportunities for empathy in our science. *Ergonomics* 44(14), pp. 1278–1289 (2001)
31. The King's Fund: Experience-based co-design. <http://www.kingsfund.org.uk/projects/ebcd> (accessed: January 17, 2015)
32. Tibben, A., Stevens, M., De Wert, G., Niermeijer, M., Van Duijn, C., Van Swieten, J.: Preparing for presymptomatic DNA testing for early onset Alzheimer's disease/cerebral haemorrhage and hereditary Pick disease. *Journal of medical genetics* 34(1), pp. 63-72 (1997)
33. Topol, Eric M.D.: *The Patient Will See You Now: The Future of Medicine is in Your Hands*. Basic Books (2015)
34. U.S. Food and Drug Administration (FDA): Warning Letter to 23andMe, Inc. (2013) <http://www.fda.gov/ICECI/EnforcementActions/WarningLetters/2013/ucm376296.htm> (accessed: January 17, 2015)
35. Udesky, L.: The ethics of direct-to-consumer genetic testing. *The Lancet* 376(9750), pp. 1377-1378 (2010)
36. Veatch, R.M.: Models for ethical medicine in a revolutionary age. *Hastings Center Report*, pp. 5–7 (1972)
37. Visser, F.S., Stappers, P.J.: Mind the face. In: *Proceedings of the 2007 conference on Designing Pleasurable Products and Interfaces (DPPI '07)*. ACM, New York (2007).
38. Wright, P., McCarthy, J.: Empathy and Experience in HCI. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 637-646. ACM, New York (2008)