



HAL
open science

An acceptability estimation and analysis methodology based on Bayesian networks

Giovanny Arbelaez Garces, Auguste Rakotondranaivo, Eric Bonjour

► **To cite this version:**

Giovanny Arbelaez Garces, Auguste Rakotondranaivo, Eric Bonjour. An acceptability estimation and analysis methodology based on Bayesian networks. *International Journal of Industrial Ergonomics*, 2016, 53, pp.245-256. 10.1016/j.ergon.2016.02.005 . hal-01522502

HAL Id: hal-01522502

<https://hal.univ-lorraine.fr/hal-01522502>

Submitted on 8 Apr 2019

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.

An acceptability estimation and analysis methodology based on Bayesian networks

Giovanny Arbelaez Garces^{a,*}, Auguste Rakotondranaivo^a, Eric Bonjour^a

^a *Université de Lorraine, ERPI, EA 3767, Nancy, F-54000, France*

Abstract

As companies are forced to conceive innovative products to stay competitive, designers face the challenge of developing products more suited to users' needs and perceptions in order to be accepted, thus reducing project risk failure. Evaluating users' acceptability has become an important research problem. Current approaches leave the acceptance evaluation question to be answered in the last stages of product development process (NPD), when an almost finished prototype is available and when there is no time left for important modifications. Acceptability evaluation methods suitable for use from the early stages of the NPD process are thus needed. This paper proposes a method for acceptability evaluation and analysis that can be used in the early stages of the development cycle. It is based on the evaluation of the solution concept by the users. The relationships among the factors (or criteria) are made explicit, thus helping designers to identify the key factors for acceptance. As the users' tests and the maturity of the concept prototype are limited in this stage, the proposed method exploits the inference properties of Bayesian networks making it possible to make useful estimations and allowing the exploration of actions that could improve the product acceptability level. Two case studies are presented in order to illustrate the method, the first related to a technological product design for a home-health care service provider and the second to a work-related musculoskeletal disorder prevention software design.

Relevance to industry: The article describes an acceptability assessment and an analysis approach to be used by industrial engineers, designers and ergonomists in the early phases of design projects. The method can help the design team to identify the levers (key factors) for enhancing product acceptance and to identify different actions (e.g. product modification, deployment strategy, and training).

Keywords: acceptability evaluation, product design, innovation, Bayesian networks, user preferences

1. Introduction

Innovation in product development projects is associated with risk and uncertainty. Besides responding to users' needs and requirements, designers should anticipate consumers' acceptance of their innovation to avoid undertaking projects that will not succeed. Acceptability evaluation is one of the key problems in product development (Luo et al., 2011).

To define acceptability as treated in this article, we should consider the technology adoption process as a temporal *continuum*. Three moments of analysis could be considered: acceptability *a priori*, acceptance, and appropriation (Terrade et al., 2009). The first one refers to the subjective representations of technology use, *i.e.* trying to predict technology usage from the perspective of what is "perceived" by the individual, even before the technology has been totally developed. In this context, dimensions such as "perceived utility" and "perceived ease of use" should be taken into account. Acceptance refers to the study of the factors which had an impact on the first interactions between the technology (developed) and the user. Finally, once the system has been deployed, it is convenient to study and assess its effective appropriation and usage.

*Corresponding author

Email address: giovanny.arbelaez-garces@univ-lorraine.fr (Giovanny Arbelaez Garces)

In the scientific literature several methods for acceptability evaluation have been proposed. Some of the approaches deal with the use of subjective and even unconscious feelings, about a product, to be translated into concrete product attributes (Luo et al., 2011; Camargo et al., 2014). The latter proposed an approach to integrate user’s perception and identify the importance and interaction of semantic attributes, thus allowing designers to validate and adjust product concepts. Another kind of approach is related to the need assessment process. The proposed methodology based on the Kano model presented in (Ben Rejeb et al., 2008), can be used as a decision aid tool for selecting consumer needs. Other approaches are based on the theory of technology acceptance, in which the users’ behavioral intention could be used as a measure of acceptability. They derive from Ajzen and Fishbein’s Theory of Reasoned Action (Ajzen & Fishbein, 1980). The most influential approach deriving from this trend is the Technology Acceptance Model (TAM) (Davis, 1989); the theory states that a user’s behavioral intention for a new technology is influenced by its perceived usefulness (PU) and its perceived ease of use (PEOU). PU was defined by the author as “the degree to which a person believes that using a particular system would enhance his or her job performance” and PEOU as “the degree to which a person believes that using a particular system would be free from effort”. The TAM model was expanded first as TAM2 (Venkatesh & Davis, 2000) to include other factors influencing the PU, and later to include factors from the organizational context of the user in the unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al., 2003). Another model (mainly used in usability studies) has been proposed by Nielsen (1993). This model distinguishes the acceptability factors of a system in two categories: social acceptability and practical acceptability. The former refers to the system’s compliance with social needs, whereas the latter is related to the technical environment of the proposed system (reliability, compatibility) as well as aspects related to its use (utility, usability). It evolved as the ISO 9241-210:2010 norm and could be used as a prescriptive tool in product design.

The models proposed in the literature lack a global vision of the influence factors that could help master its complexity. Furthermore, those models are mostly used as measure mechanisms once the technology has been deployed thus requiring a finished prototype and a huge amount of users’ test. Consequently, this paper aims to present a method for acceptability evaluation that can be used to guide the design process, helping designers to identify the levers (key factors) for acceptability from the early stages of the design process. As prototypes and users’ test data are scarce in the first stages of product development, some artificial intelligence techniques can help deal with those drawbacks. There have been some propositions to evaluate and predict acceptability using some of those techniques. Elazouini et al. (2005) for example, estimated the acceptability of a new formwork system for construction using neuronal networks; Luo et al. (2011) proposed a combination of techniques (Radial Basis Function Networks, Support Vector Machine Sequential Minimal Optimization and their ensembles) along with Bayesian networks to predict acceptability of products. They tested the methodology with a car evaluation; their methodology, however, requires large data-sets. Compared to other techniques (neural networks, support vector machines, decision trees) a Bayesian network allows the integration of different sources of information (experts’ knowledge, and data), allows the handling of uncertainty and incomplete data and provides a visual graph of criteria. Therefore, the basis of our proposed method resides on the usage of Bayesian networks and their inference properties.

The remainder of the paper is structured as follows: in Section 2 we present an overview of the properties related to Bayesian networks. Section 3 describes the proposed methodology. Section 4 presents the case example. Results and implications for theory and practice are discussed in section 5. Finally, conclusions and perspectives for future research are presented in section 6.

2. Reminder on Bayesian networks

Bayesian networks are directed acyclic graphs that represent the probabilistic relationships between a group of variables (Pearl, 1988). They are generally composed of two parts: a quantitative part and a qualitative part. The qualitative part is the graph in which the nodes represent the states of random variables and the arrows pointing from a parent node to a child node represent the causal conditional dependency between the two nodes. As for the quantitative part, it mainly refers to the relationships between a node and its parents. The relationship can be represented by the probability that the node’s state will provide

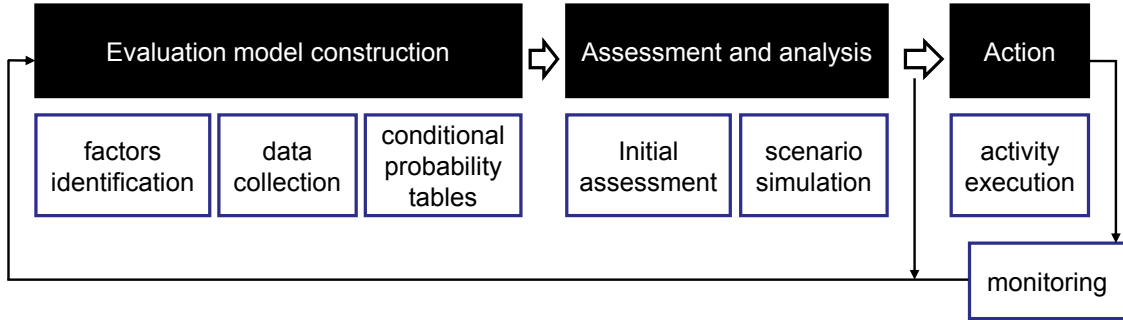


Figure 1: Overview of the methodology.

different probabilities of the parent’s node state. This information is consigned to the conditional probability table (CPT) of the node.

Two types of reasoning are possible with Bayesian networks: (1) diagnosis (backwards inference) which, given an observation, makes it possible to find the most probable cause among the hypotheses, and (2) prediction (top- down inference), which makes it possible to estimate the probability of an observation given the assumptions (Jensen & Nielsen, 2007).

In the framework of our research work, these properties are interesting since they will allow us to: calculate and predict the level of acceptability of a new technology (acceptability index), to identify among the criteria the key levers for acceptability improvement and to estimate the impact of different scenarios of action on the computed acceptability index. The aim is to reduce the risk of project failure by improving the product acceptability. Indeed, Bayesian networks have been shown to be successfully used in risk management. For instance, Hu et al. (2013) used them in the analysis of software projects risk; and Song et al. (2013) used them to assess the risk of service failures by focusing on ripple effects.

3. Proposed methodology

The objective of the proposed methodology is to contribute to the concept validation process by estimating the acceptability index of a solution concept and helping the designers to identify how factors contribute to this index. Once the initial estimation is made, the proposed approach helps in the analysis of different sets of actions (scenarios) that are most suited to improve the users’ acceptability of the solution concept. The proposed method is composed of three main phases (Figure 1): evaluation model construction, assessment and analysis and finally, action.

3.1. Evaluation model construction

The aim of this phase is to build the acceptability evaluation model. The first step consists in the identification of the acceptability criteria. The second step consists in collecting data about those criteria (or factors) by means of user evaluations of the solution concept. The third step uses the collected data to build a Bayesian network (BN) for the acceptability assessment. The BN is build by generating the conditional probability tables (CPT) of the network from the collected data.

3.1.1. Factors identification

We propose to use a user acceptance model like TAM or UTAUT (Davis, 1989; Venkatesh & Davis, 2000; Venkatesh et al., 2003) to define the factors that should be evaluated. To facilitate model development, factors can be divided into two groups: those related to the user (*e.g.* perception of utility, perception of ease of use) and those related to the context of usage (*e.g.* social influence, facilitating conditions). Analysts can review the literature of the domain concerned to find related factors. As an example for a new medical diagnosis application, in the TAM model two criteria explain the intention to use: perceived ease of use and perceived usefulness (utility). One can define composing factors to evaluate each one of

those criteria. The utility of the medical application relies on its diagnosis and accuracy capability, and so perceived usefulness can be composed by “diagnosis is accurate” (users’ perception of device diagnosis accuracy) and other factors. A group of experts should be consulted to facilitate, complete and validate the evaluation model.

3.1.2. Data collection

In this step a questionnaire is prepared, including all the criteria and their composing factors identified in the previous step. Participants in the test are asked to evaluate the prototype as regards each factor on a pre-defined scale. The 5- to 7-level Likert scale with extremes from “totally disagree” to “completely agree” can be used (Sinclair, 2005; Likert, 1932). As in the user acceptance questionnaires, a global acceptability item (intention towards use) must be included in the evaluation to serve as a final utility score. Depending on the items defined in the previous step, the questions can be complemented using physiological measurements like electromyography (EMG), galvanic skin response (GSR), electroencephalography (EEG), heart rate (HR), temperature, and respiration rate for example (Silva et al., 2013).

3.1.3. Generation of the conditional probability tables

Once data from the test have been collected the analyst can build a Bayesian network to estimate the acceptability level. Any software that allows Bayesian networks modeling ¹ can be used in this step. The data are manually exported to the software, each factor is represented as a node, and the links between them (arrows) are fixed as previously defined in the factor identification step according to the acceptability model. In the export process the probability distributions for each variable (CPT) are built from the questionnaire data using a parameter learning algorithm (Dempster et al., 1977). The information contained in the CPT reflects the weights and contributions of each criterion to the acceptability index. An alternative to this step is to load the nodes from the test data and use a structure learning algorithm to automatically build the structure of the network (Cooper & Herskovits, 1992; Oniško et al., 2001). Naive Bayes learning using the “acceptability” node as target is a good starting point; the missing links or incorrect links need to be modified by the analyst and validated by a group of experts (Domingos & Pazzani, 1997).

3.2. Assessment and analysis

Once the Bayesian network is built, the initial assessment of the acceptability level, given the obtained data, is made (the inference algorithm (Huang & Darwiche, 1996) is used). With the initial assessment results the analysis can start, different value propagations can be made to each factor to compute its profile at different states. Designers can use those simulations to analyze different scenarios and find among the factors those that contribute the most to the global acceptability (key levers for acceptance). For example, in a medical diagnosis application, a characteristic evaluated by users could be that “diagnosis is accurate”. Application designers can compute the users’ response as if they have perceived either to its maximum or minimum levels (1 not accurate and 7 highly accurate for example). The value of acceptability is observed for each change (what happens to the global acceptability when “diagnosis is accurate” is evaluated at 1? and at 7?). If this factor seems to highly influence the “intention to use” when evaluated positively by the users (the acceptability index shows a high value (> 6); when “diagnosis is accurate” takes 7 value), then designers should consider how to improve this value; that is, make the diagnosis of the application more accurate, for example by improving the diagnosis algorithm.

3.3. Action and Monitoring

Action step corresponds to putting into practice the actions or improvements that have been found and selected through the simulations in the previous step. In this phase, the design team will try to improve the factors that have been proved to have a significant impact on the acceptability of the product (diagnosis accuracy for the example). Our intention in adding this step is that once the action has been performed a

¹In this research, we used the proprietary BayesiaLab (www.bayesia.com); there are several others freely available, GeNIe, Netica, R are some examples.

new test (e.g. survey, user test) could be made and the results could be updated in the network. In this way, it makes it possible to have a new estimation of the acceptability level, thus allowing the changes made to be monitored. The latter activity corresponds well to the iterative and progressive nature of the design process. This step will not be developed in the real case studies described in this paper.

4. Application of the methodology

The methodology was applied in two real cases. They are presented in the next section. The cases differ in the level of expertise of the users and in the level of maturity of the prototype used. The first case uses the previous experience of the users over a particular solution to tackle the design problem with a proposed solution concept. The second case starts with a design concept proposed to prospective users without previous experience on similar products.

4.1. First case description: HADAN's communicating pen

The HADAN is the in-home healthcare network of the agglomeration of Nancy (France). It is a not-for-profit association in charge of coordinating in-home healthcare and in-home hospital care for patients and linking several structures (e.g. hospitals, physicians, nurses, and other health care professionals). In 2008 an innovative communicating pen was introduced to the structure (Augustin et al., 2010). The pen was designed to enable nurses and physicians to instantly send a set of actions performed in a patient's home to a central server and thus improve care practices (e.g. traceability, reducing time, coordination, information sharing, instant billing...). The care practitioners had to fill out a paper form with the information required; the pen registers all the information filled out. If they tick the *send* case and put the pen in the charging socket, the information is sent through a GPRS connection to the central server. Then it is stocked and made available for consultation by the interested stakeholders (physician, coordinator, pharmacist and nurses). Whereas the evident utility of this device was recognized by the personnel, its introduction to the structure was not easy. There were problems with the information not being properly recognized and stocked in the database. With the intention of providing a better device a graduate-student from our research group conducted an analysis from the pen experience and an evaluation of user's acceptance (using TAM model). Using this information she proposed a new concept device (Esparza Ribera et al., 2011). The data collected from this analysis were used to illustrate the following use case scenario for the proposed method for acceptability evaluation. The conceptual design proposed is a tablet-like device with an interface regrouping all the forms needed by practitioners. Information is transmitted (through a wireless connection) to a central server and shared between all involved actors (Esparza Ribera et al., 2011). This solution requires a big shift in practices as paper forms are no longer required; all information has to be submitted through the tablet application.

4.1.1. Evaluation model construction

In the first case our aim was to test if the methodology was applicable. We selected the TAM model due to its simplicity and because of its power to explain the acceptability of a technology with two main criteria.

Factors identification

In the TAM model there are two groups of criteria that explains the intention to use a technology (ITU), those related to the perception of ease of use (PEOU) and those related to the perceived utility (PU) (Davis, 1989). Each one of those criteria was completed by analyzing and decomposing them on factors to be evaluated. The utility of the device is in helping nurses and physicians to reduce the medical information processing time (PU2) by improving communication of the performed acts between the different parties (e.g. pharmacist, other nurses, coordinators) and by automating the billing process. It is also useful if the transmitted messages and information contained in the forms are of good quality (PU3) and finally, the device is useful if it works as intended (PU1), by not generating wrong messages for example. To evaluate the perceived ease of use it is necessary that users evaluate the facility to use the device (PEOU1), the ease required to complete the forms (PEOU2), if the ergonomics of the device are appropriate (to operate, to handle for example) (PEOU3) and finally if the forms are adapted to their required use (PEOU4). The

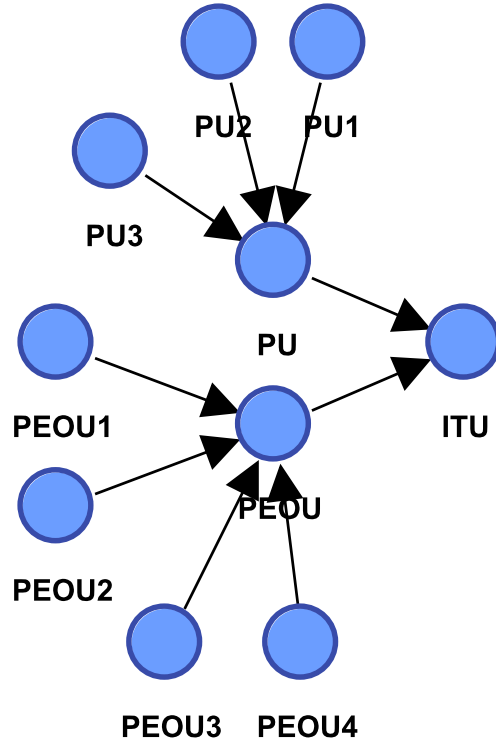


Figure 2: Acceptability model for the communicating pen.

model in Figure 2 shows the acceptability modeled by the users' intention to continue to use the device with the variable (ITU) with the associated identified factors. Table 1 shows the description of the nodes.

Data collection

A survey based on the previous model was constructed to evaluate the users' perception of the product. We selected among the active users of the communicating pen 18 participants to answer the questionnaire. They were selected to represent the three main groups of users: 6 managers, 5 physicians and 7 nurses. They were asked to evaluate the device using a 5-point Likert scale from "not at all" to "very much" for each factor and to base their answers in their perception of the potential usage of the device (Sinclair, 2005).

Collected data were verified to evaluate consistency. We tested the normal distributions of data with a Kolmogorov-Smirnov test (see Table A.3).

Generation of conditional probability tables

The model and factors identified in the first phase were modeled as a Bayesian network using the BayesiaLab software. Each network node corresponds to each identified factor. The survey answers were associated to the modeled network and the conjoint probability tables of each node were computed using the EM algorithm (Dempster et al., 1977). The modeled network was evaluated to verify its performance given the data; the total precision of the network is 94.4% (function of the correct predictions of the target variable) with $r = 0.962$.

4.1.2. Assessment and analysis

With the evaluation model created in the software, the initial acceptability assessment is made. Figure 3 shows the distribution of the data.

From the data we can observe that users were willing to continue the experience with the device (ITU mean: 4.222 in a scale of 5) and that they perceived the utility of the device as medium (PU mean: 2.883).

Table 1: Code use for the model nodes.

Code	Factor description
PU	Perceived utility
PU1	The device functions as intended
PU2	The device helps to gain time during the procedures
PU3	The transmitted messages are of good quality
PEOU	Perceived ease of use
PEOU1	The device is easy to use
PEOU2	The forms are easy to fill out
PEOU3	The device ergonomics are appropriate
PEOU4	The forms are adapted to required usage
ITU	Intention towards use

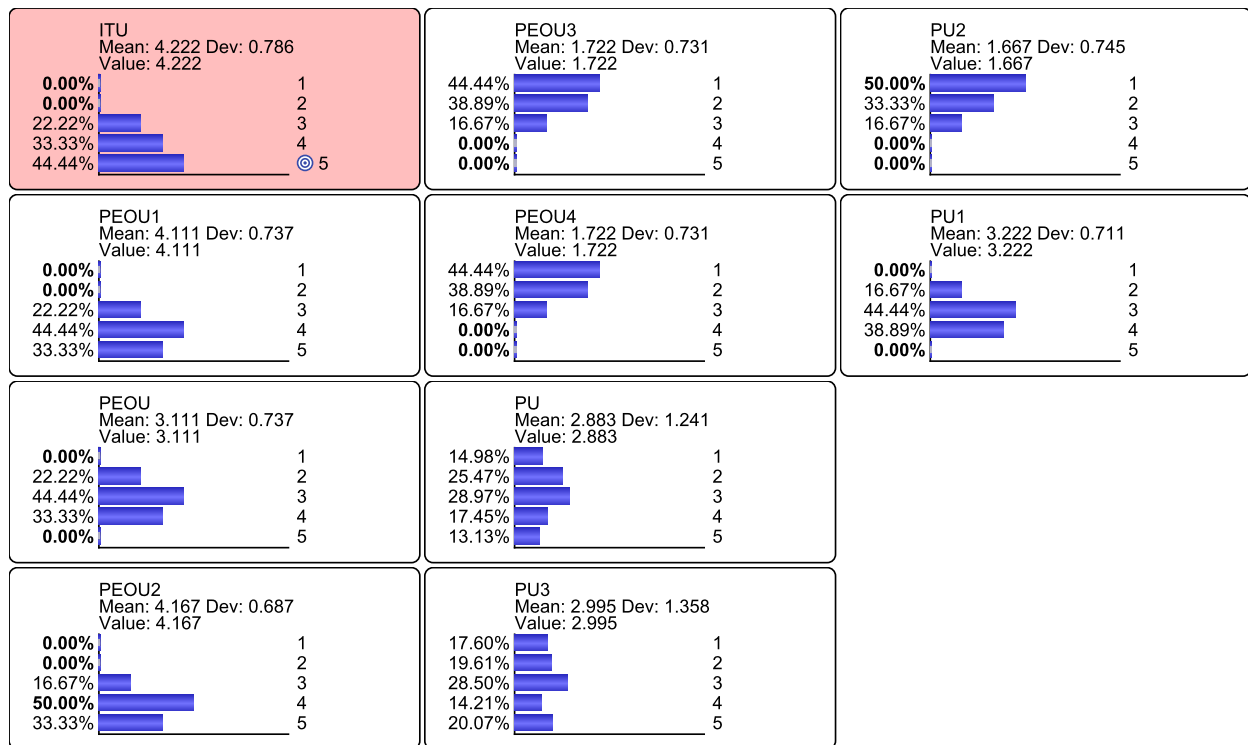


Figure 3: Acceptability estimation for the communicating pen.

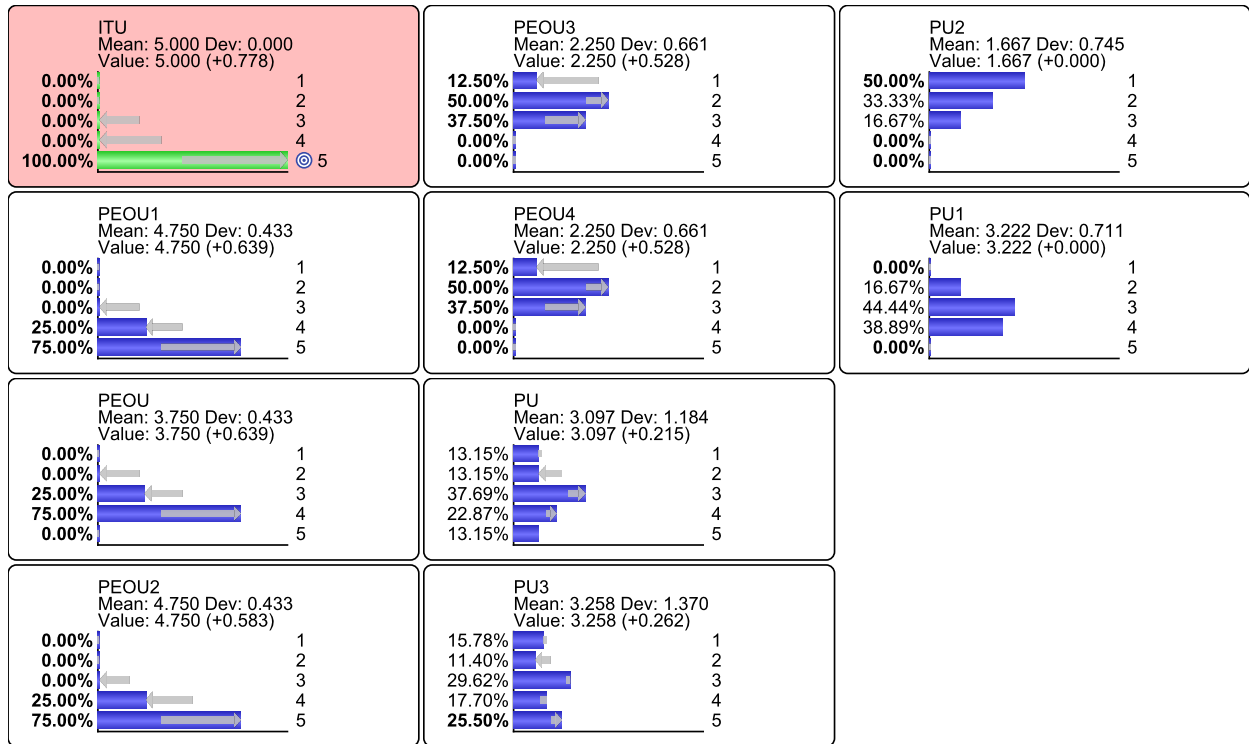


Figure 4: Intention to use (ITU) value 5 profile.

As previously stated in the model, we defined that the users' intention to use the device would be influenced by two groups of variables (PU and PEOU). Once the initial assessment had been made, the design team had to look for those factors that are key levers of acceptability. They had to look for actions to improve the users' perception concerning the factors and then to increase the acceptability index. This is done through the simulation and analysis of different sets of actions (scenarios).

Figure 4 shows the respondents profile with a high intention to use (ITU=5). All the variables change showing what the profile of this type of respondent is more likely to be. The highest variation (+0.639) is concentrated in the perceived ease of use nodes (PEOU1, PEOU2, PEOU). This means that users that perceive those factors with a high-level are more likely to use the device (ITU=5). In other words, they are more likely to accept the device. The product design team should seek to improve these factors by working on: facilitating the form filling and improving the ease of use of the device. The ease of use could be improved by reducing the steps required to send the information or by increasing users' training.

In Figure 3 we can see that the messages' quality (PU3) is perceived with a mean of 2.995 ($sd = 1.358$). In the experience and comments collected from the users, it was found that sometimes the ADELI number (code required to treat the patient's information) was not properly recognized when transmitted to the database. This is reflected by the users' perception of a poor message quality even if no problem was reported for the other information transmitted. The design team can then simulate the effect of improving message quality. Figure 5 shows the profile of the network when the messages' quality (PU3) takes the value of 5. From the computation we can observe that the node of interest ITU takes a high value (4.380) reflecting that the proposed changes would improve product acceptability. The improvement action could be included in the project planning to be implemented.

4.2. Second case description: Active pause software

The second case presented corresponds to a software application design and development for prevention of the risk of repetitive strain injury and other work-related musculoskeletal disorders (WMSDs) when

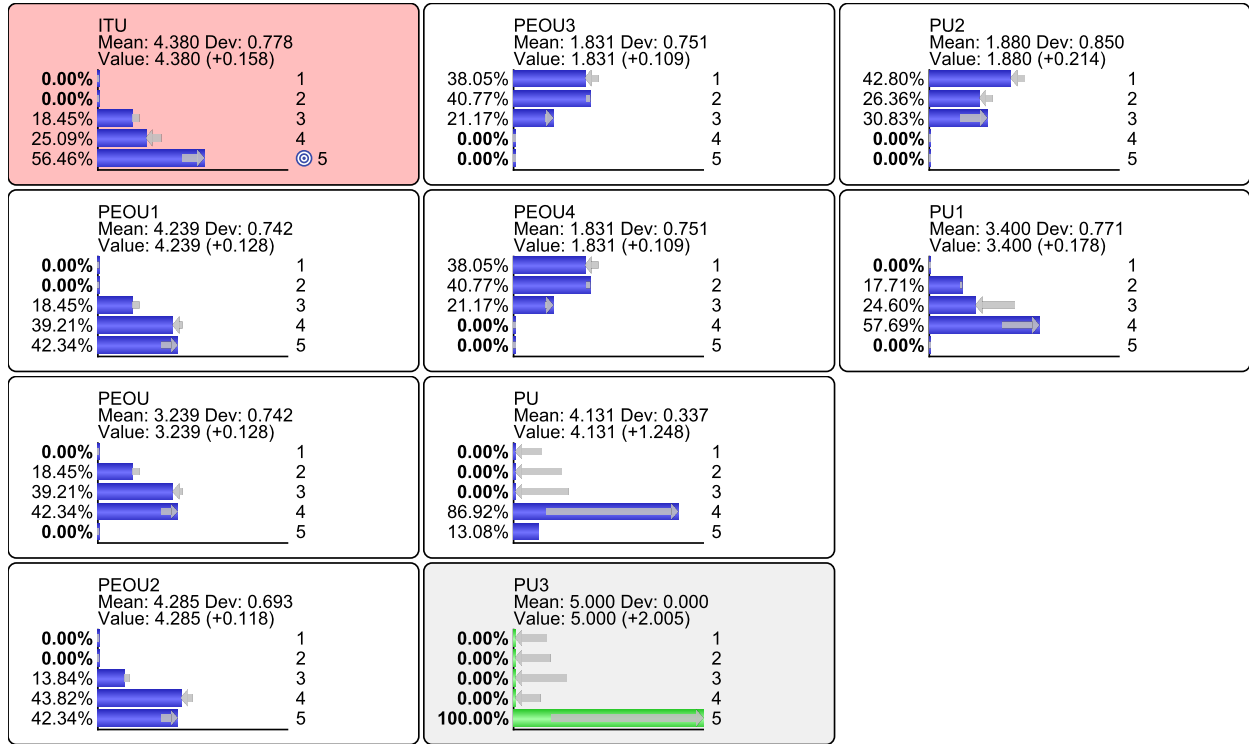


Figure 5: Network profile for a high message quality (PU3=5).

frequent computer work is done. Different studies have found a correlation between long computer usage and musculoskeletal problems (Chang et al., 2007; Cho et al., 2012; Eijkelhof et al., 2014). In France, the Institute for Research and Security (Institut national de recherche et de sécurité, INRS) advises that computer workers are permitted frequent pauses to reduce fatigue (INRS, 2011). Studies support that break reminder software and exercise during the brake (active pause) can help prevent those injury risks (Irmak et al., 2012; Henning et al., 1996). For this reason, a proposal for active pause (break reminder and RSI prevention exercise) software is made. As in the early phase of the product design a functional prototype is not always available, the stimuli used to evaluate users' acceptability and perceptions in this case study is based on a low fidelity prototype (Lim et al., 2008).

4.2.1. Evaluation model construction

In the first case study we verified that the methodology was applicable. Therefore, the acceptability model used was simple. We notice that the factors related to the organizational context were not taken into account. Aspects related to the perceived influence of the colleagues or superiors for example were not considered. For this reason, in the second case we decided to use an acceptability model that includes the socio-organizational aspects.

Factors identification

The model for evaluating acceptability was constructed using the TAM model as reference. In this case study, the factors of social influence and facilitating conditions related to the context of usage, as presented in the UTAUT theory, were also taken into account (Venkatesh et al., 2003). The perceived utility (PU) of the software was decomposed by the factors: its ability to help to reduce repetitive strain injury risk (PU1) and to help to minimize fatigue during prolonged computer work (PU2) and the users' appreciation of the software usefulness. The global Perceived ease of use (PEOU) is evaluated by the users' perception of the software use without help (PEOU2) and their perception of the ease of use of the software (PEOU1). The

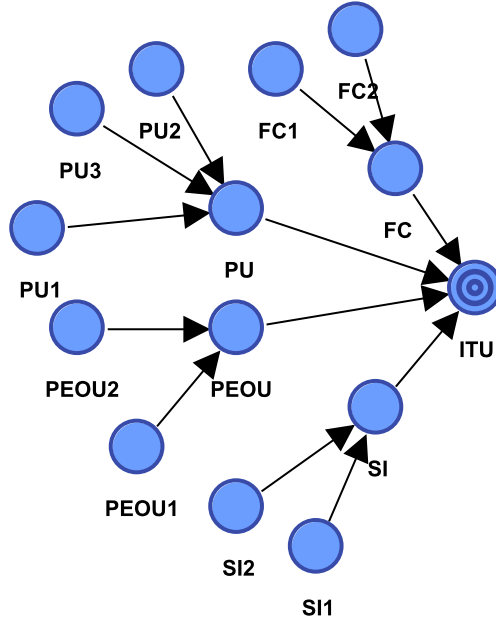


Figure 6: Acceptability model for the Active pause software.

organizational facilitating conditions (FC) were evaluated by the users’ perception of being able to use the software at their workplace (FC1) and their perception of the control they will have over the software (FC2) (i.e. will the users be able to activate or deactivate the software, will it disturb their work during meetings for example). The perceived social influence (SI) was evaluated by the perception of their colleagues (SI1) and their superiors (SI2) about the software used in the workplace. The model in Figure 6 shows the acceptability model developed to evaluate this product and table 2 shows the description of the nodes.

Data collection

A questionnaire adapted to evaluate user’s product perception was build. The survey had nine questions plus one concluding question on the intention to use the proposed software. We used a 7-point Likert scale ranging from “not at all” to “very much” (Sinclair, 2005).

We showed the survey respondents a short text introducing the solution and a description of the software functionality (break reminder according to one’s work, exercise propositions to prevent RSI and other WMSDs). A screen capture of the proposed software interface was also shown. The image showed an exercise for the prevention of eye fatigue and the duration of the break.

Faced with this stimulus, users were asked to respond to the questionnaire. The workers in academic and research environments (e.g. researchers, contractual graduate students, teachers and administrators) are likely to be subject to these professional risks because the usually spend more than 2 hours of continued computer work each day. We decided to explore the prospective opportunity of this type of software among the workers of an engineering school in Nancy, France (about 51 persons involved). An email was sent inviting them to take the survey. We obtained 28 answers from the prospective users (15 female; 13 male; age range: 18-50 years). From the group surveyed 14 % spend between 2-4 hours per day using the computer, 45% spend between 4-8 hours and 41% spend more than 8 hours. Collected data were verified to evaluate consistency. We tested the normal distributions of data with a Kolmogorov-Smirnov test (see Table A.4).

Generation of conditional probability tables

The Bayesian network was constructed in the BayesiaLab software according to the model presented in Figure 6. The collected data from the survey respondents was associated to the network. The conditional probability tables were generated using the parameter learning algorithm proposed by Dempster et al.

Table 2: Code use for the model variables.

Code	Factor
PU	Perceived utility
PU1	The software will help to reduce repetitive strain injury (RSI) risk
PU2	The software will help to minimize fatigue at work
PU3	I perceive the software as useful
PEOU	Perceived ease of use
PEOU1	I perceive the software as easy to use
PEOU2	I will be able to use it without help
FC	Facilitating conditions
FC1	I will be able to use it at my workplace
FC2	I will have control over the software (i.e. stop or pause during a meeting)
SI	Social influence
SI1	My colleagues will agree with software use
SI2	My superiors will agree with software use
ITU	Intention towards use

(1977). The conditional relationships represent the contribution of each factor in the model to the ITU. As we are going to observe in the next step those weightings are different for each factor given that some have more influence in the final acceptability. The modeled network was evaluated; the global precision is 78.57% (function of the correct predictions of the target variable) $r = 0.847$.

4.2.2. Assessment and analysis

Figure 7 shows the initial acceptability assessment according to the survey data.

We observe that, concerning the users' perceived intention to use (ITU) the proposed active pause software has a mean value of 4.6. The graphical distribution of the results shows that 67.86 % of the users' intention to use (ITU) is in the upper high-level (5, 6, 7). However, those values alone do not give designers much useful information. To look for the factors that have a high influence on users' intention to use, the designers follow the simulation approach. They look for the profile of the ITU index when it takes the maximum value (7)(see Figure 8).

From this inference we observe that the users' perceived utility (PU3) moves from an average of 4.786 to 5.667. From the model constructed, we know that this factor is related to the users' perception that the software could help reduce repetitive strain injury risk (PU1) and that it will help to minimize fatigue at work (PU2). The former factor varied from a mean of 4.929 to 5.542 and the latter from a mean of 4.964 to 5.527.

The two factors related to social influence (colleagues' influence towards use (SI1) and superiors' influence towards use (SI2)) also have an impact on the value of the users' intention to use (ITU) the software. A simulation for the social influence (SI) on level 7 is made. In Figure 9 we can observe that the users' intention to use (ITU) moves to a high-level (ITU mean = 5.804 and 60.71% respondents in level 7). This means that users perceiving a positive influence of colleagues and superiors are more likely to have a positive influence towards using the software. This may also imply that users perceiving their colleagues' refusal to use the software would be more likely not to use the software themselves.

The design team can repeat and simulate as many scenarios as they can imagine with the data available. With the two observations previously made, two courses of action can be imagined to increase users' acceptability (intention to use). Actions intended to increase the perceived utility of the application and actions to improve the perceived social influence. The first kind of action, could be focused on the product design perspective, i.e. improvements made to the software to increase the utility, for instance, to emphasize the exercises that users can do to reduce work-related musculoskeletal disorders risk presented by the application. In addition, they can also have a marketing or organizational change perspective, i.e. software deployment can be accompanied by training or advertising campaigns to raise users' awareness of the benefits of the

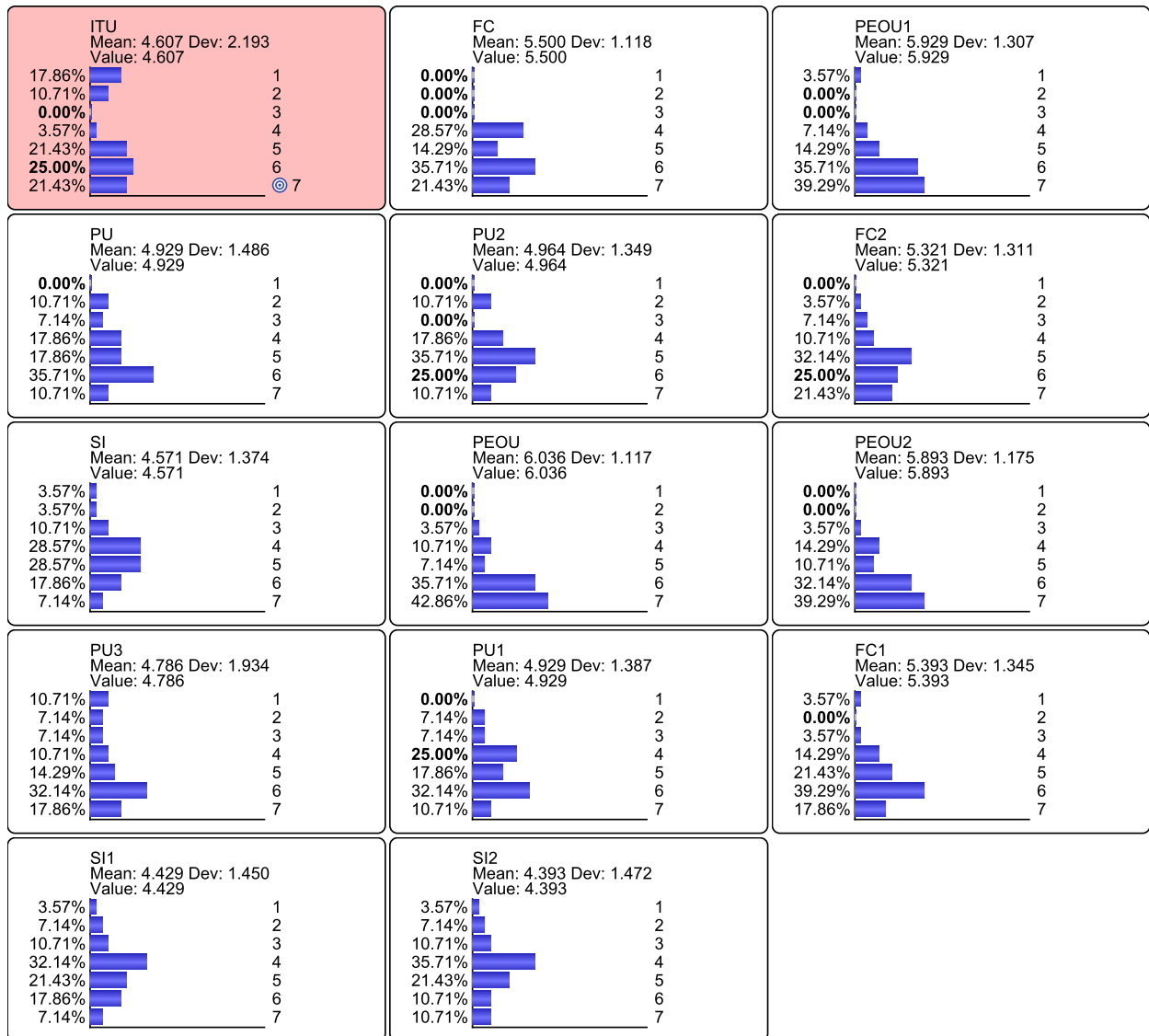


Figure 7: Acceptability estimation for the active pause software.

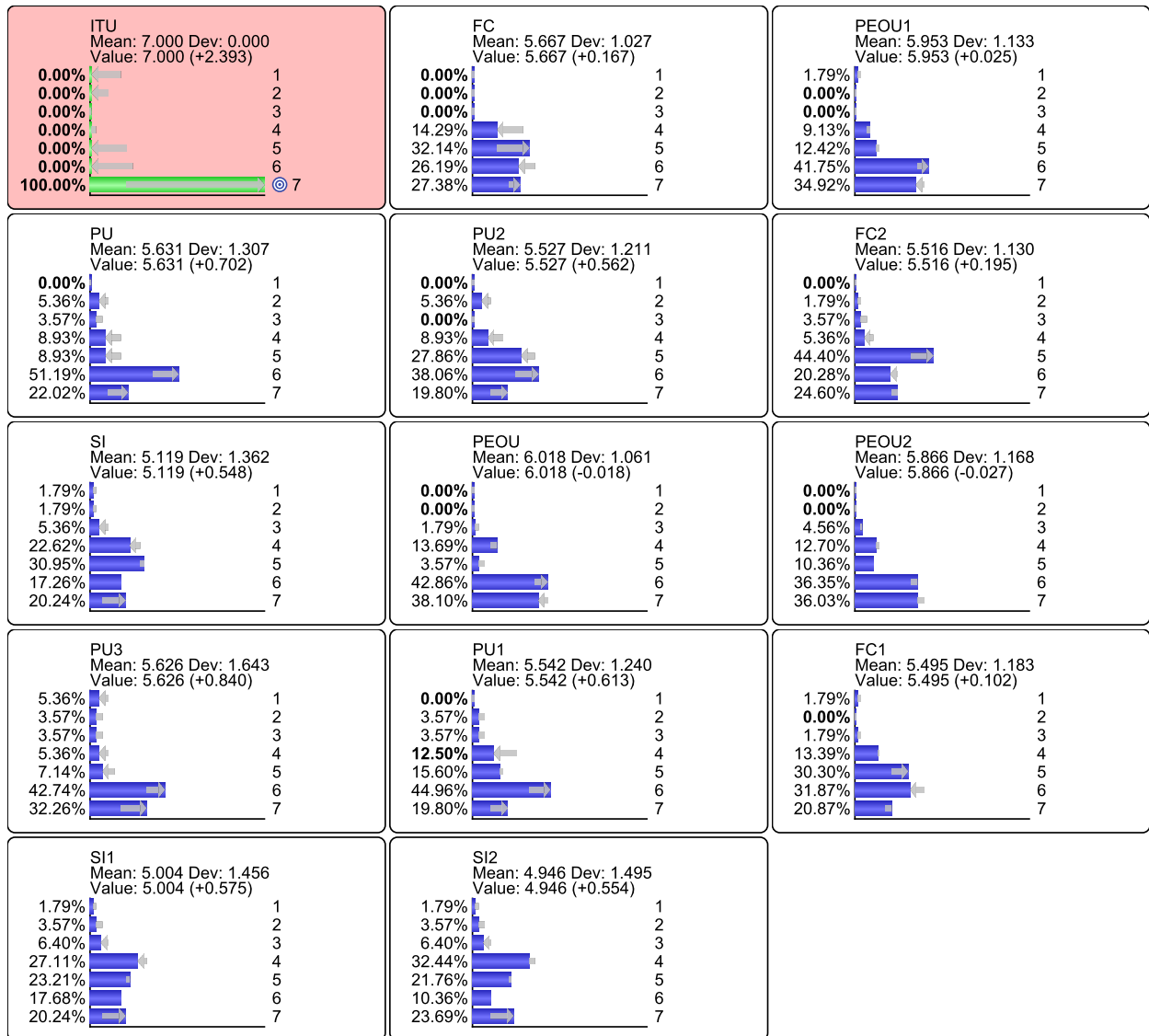


Figure 8: Users' intention to use (ITU) profile for the maximum value (7).

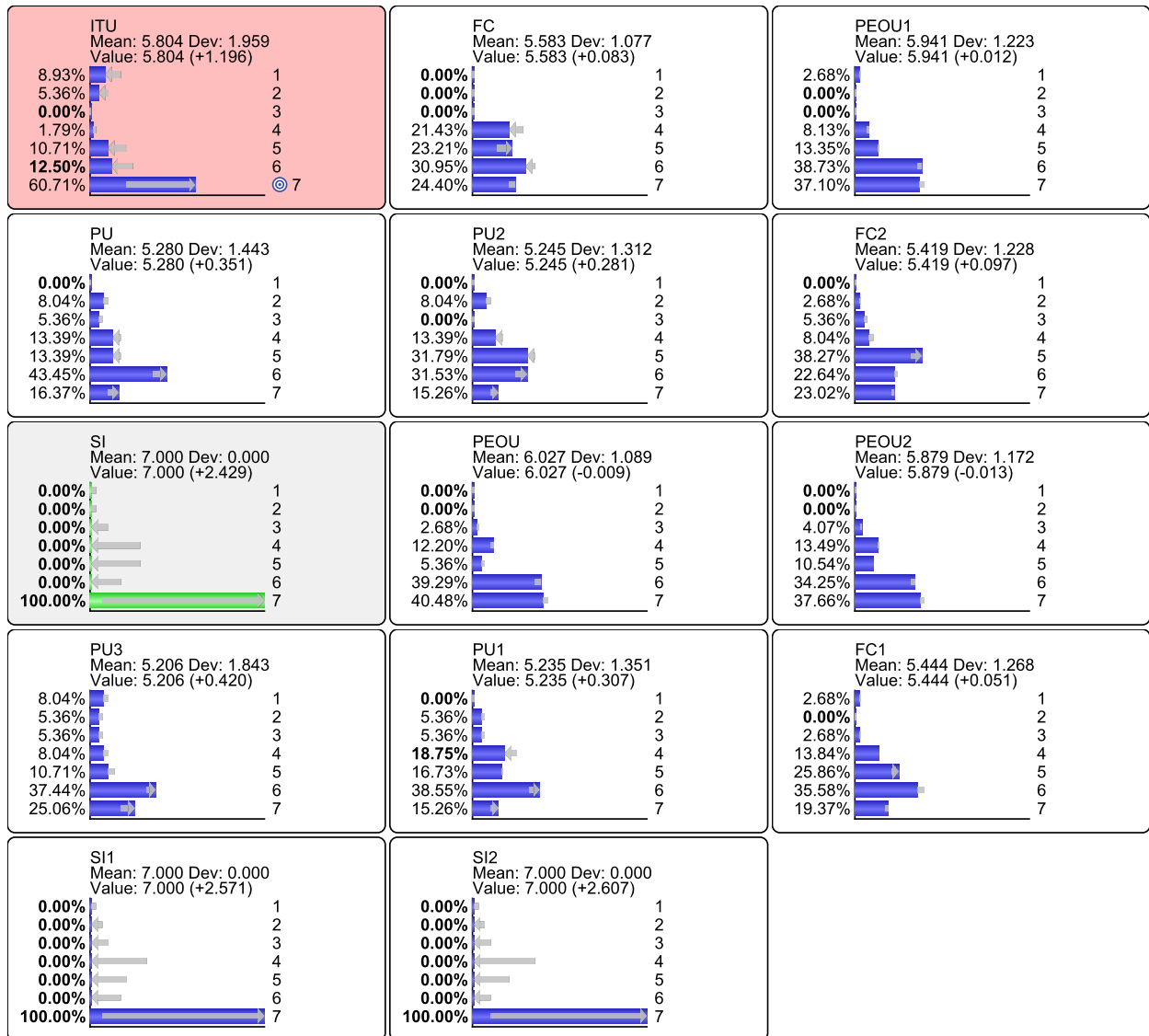


Figure 9: Social influence to use (SI) profile for the value seven.

application; they can provide statistics or scientific reports related to workers suffering from work-related musculoskeletal disorders by not practising the appropriate preventive actions for example. The second kind of actions, aiming at improving the social influence aspect, concerns the organizational change and marketing actions, for instance, campaigns or actions that follow deployment. For example, superiors can be the spokesperson for software deployment in the company. Also, they can select a few workers (key users) to be pilot users that later become the reference for convincing and encouraging other colleagues to use the software ([Berchet & Habchi, 2005](#); [Moitra, 1998](#)).

5. Discussion

The proposed methodology for acceptability evaluation and analysis completes existing approaches estimating the acceptability of a new concept. The methodology allows designers to identify and thus master the complexity of the factors which have a strong influence on final product acceptance. The identified relationships between factors and acceptability allow designers to analyze and look for strategies to improve them. The proposed method exploits the properties of Bayesian networks for knowledge modeling and inference in order to deal with the uncertainty and the lack of users' test information on this early phase. Two design case applications are used to illustrate the proposed methodology. When compared to the [Luo et al. \(2011\)](#)'s proposal, the method we propose in this paper differs as it is based on few data; its interest resides on having user tests with few users and this with the aim of quickly modifying the stimuli or prototype used and then testing once more with the users. However, in a domain where historical users' data are available, it can be used to construct the first Bayesian network and take advantage of the accuracy proposed by the [Luo et al. \(2011\)](#) method to start the first phase of testing. We also want to make explicit the relationships among the different factors as in the proposal of [Camargo et al. \(2014\)](#). We also provide a way to estimate the effects of changes in those relationships allowing the design team to analyze different scenarios of action before putting them into practice.

Even if the details and design implications of the two case studies are not the main purpose of this article, the information provided can help to better understand users' preferences. As already shown, the design implications and strategies for acceptability improvement do not concern the design phases only. Project managers and the entire product development team can modify the development/deployment strategies according to the findings. Design decisions can find a compromise between product modifications and deployment strategies (for instance, training for users, for maintenance agents). The team could choose not to modify the product design but to work in the deployment strategy and communication.

One limit of the proposed method is related to the dependence on the users' test and its design. For example, in the first case presented users had previous experience with a similar product. So they could easily project themselves into the usage situation and thus we can place more confidence on their answers. This is somewhat similar to concept testing in a market research evaluation and is suitable for incremental innovation projects. In the second case users had no or limited experience with similar products, and the stimulus used in the concept evaluation was a low fidelity prototype. This situation has some similarity with future studies in market research which might be the case of radical innovations. [Sauer & Sonderegger \(2009\)](#) and [Sauer et al. \(2010\)](#) have studied the relationship between users' experience and prototype fidelity on usability testing. While they show that low fidelity prototypes are generally suitable for predicting user usability, in the case of radical innovations such as the one shown in the second case it might require a mixture of appropriate strategies between users' experience and prototype fidelity levels in order to ensure that users are able to represent themselves in the future usage situation ([Roland Ortt et al., 2007](#)). Further research might explore different strategies for user testing; we can consider for example the usage of web-based surveys as proposed by [Cho et al. \(2011\)](#).

One of the drawbacks of the current models for acceptability evaluation is that they only take into account the perception of one stakeholder: the user, while in a social context the other stakeholders should be taken into account. In its current state, the methodology we proposed in this paper makes it possible to support the collection of information from different stakeholders. The model built in the first phase can include criteria for evaluating their perceptions of the future product.

6. Conclusion

This paper presents a new method for acceptability evaluation that can be used in the first stages of product development process, particularly in the concept evaluation stage. The proposed method allows designers to identify and thus master the complexity of the factors which have a strong influence on the final product acceptance. The identified relationships between factors and acceptability allow designers to analyze and look for strategies to improve them. The proposed method exploits the properties of Bayesian networks for knowledge modeling and inference in order to deal with the lack of data in the early design phase. Two real design cases are used to illustrate the interest of the proposed method.

Another contribution of this method is that it unveils hidden relationships between the evaluated factors and acceptability. The importance (weight) of each relationship is also represented. This last feature is important in radical innovation projects where those weights are rarely available and hardly estimated by prior knowledge or experts' consultation.

In addition to the further research identified in the discussion section our current work is focused on providing a tool for the optimal evaluation and selection of the sets of actions either by fixing a target level of acceptability and optimizing the supplementary cost or by fixing an improvement budget and optimizing the acceptability level.

Acknowledgments

This work is financed by a grant from the French Ministry of Higher Education and Research (Ministère de l'Enseignement supérieur et de la Recherche, MESR).

References

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Augustin, M.-D., Rakotondranaivo, A., Grandhaye, J.-P. P., & Vuillemin, N. (2010). Performance et innovation : apports de la comptabilité analytique et du stylo communicant dans le cadre de l'hospitalisation à domicile de Nancy. In *5ème Conférence Francophone en Gestion et Ingénierie des Systèmes Hospitaliers GISEH'2010* 1. Clermont-Ferrand, 2 - 4 septembre.
- Ben Rejeb, H., Boly, V., & Morel-Guimaraes, L. (2008). A New Methodology Based on Kano Model for the Evaluation of a New Product Acceptability during the Front-End Phases. In *32nd Annual IEEE International Computer Software and Applications Conference* (pp. 619–624). IEEE. doi:10.1109/COMPSAC.2008.94.
- Berchet, C., & Habchi, G. (2005). The implementation and deployment of an ERP system: An industrial case study. *Computers in Industry*, *56*, 588–605. doi:10.1016/j.compind.2005.02.009.
- Camargo, M., Wendling, L., & Bonjour, E. (2014). A fuzzy integral based methodology to elicit semantic spaces in usability tests. *International Journal of Industrial Ergonomics*, *44*, 11–17. doi:10.1016/j.ergon.2013.08.007.
- Chang, C.-H. J., Amick, B. C., Menendez, C. C., Katz, J. N., Johnson, P. W., Robertson, M., & Dennerlein, J. T. (2007). Daily computer usage correlated with undergraduate students' musculoskeletal symptoms. *American journal of industrial medicine*, *50*, 481–8. doi:10.1002/ajim.20461.
- Cho, C.-Y., Hwang, Y.-S., & Cherng, R.-J. (2012). Musculoskeletal symptoms and associated risk factors among office workers with high workload computer use. *Journal of manipulative and physiological therapeutics*, *35*, 534–40. doi:10.1016/j.jmpt.2012.07.004.
- Cho, Y., Park, J., Han, S. H., & Kang, S. (2011). Development of a web-based survey system for evaluating affective satisfaction. *International Journal of Industrial Ergonomics*, *41*, 247–254. doi:10.1016/j.ergon.2011.01.009.
- Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, *9*, 309–347. doi:10.1007/BF00994110.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, *13*, 319.
- Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society.*, *39*, 1–38.
- Domingos, P., & Pazzani, M. (1997). On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. *Machine Learning*, *29*, 103–130. doi:10.1023/A:1007413511361.
- Eijkelhof, B. H. W., Huysmans, M. A., Blatter, B. M., Leider, P. C., Johnson, P. W., van Dieën, J. H., Dennerlein, J. T., & van der Beek, A. J. (2014). Office workers' computer use patterns are associated with workplace stressors. *Applied ergonomics*, *45*, 1660–1667. doi:10.1016/j.apergo.2014.05.013.
- Elazouni, A. M., Ali, A. E., & Abdel-Razek, R. H. (2005). Estimating the Acceptability of New Formwork Systems Using Neural Networks. *Journal of Construction Engineering and Management*, *131*, 33–41. doi:10.1061/(ASCE)0733-9364(2005)131:1(33).

- Esparza Ribera, E., Camargo, M., Grandhaye, J. P., & Rakotonranaivo, A. (2011). *Evaluation de l'acceptabilité d'une nouvelle technologie dans le domaine médicale: une analyse à l'aide des réseaux bayésiens*. Master's thesis ENSGSI Nancy, France.
- Henning, R. A., Callaghan, E. A., Ortega, A. M., Kissel, G. V., Guttman, J. I., & Braun, H. A. (1996). Continuous feedback to promote self-management of rest breaks during computer use. *International Journal of Industrial Ergonomics*, *18*, 71–82. doi:10.1016/0169-8141(95)00032-1.
- Hu, Y., Zhang, X., Ngai, E., Cai, R., & Liu, M. (2013). Software project risk analysis using Bayesian networks with causality constraints. *Decision Support Systems*, *56*, 439–449. doi:10.1016/j.dss.2012.11.001.
- Huang, C., & Darwiche, A. (1996). Inference in belief networks: A procedural guide. doi:10.1016/S0888-613X(96)00069-2.
- INRS (2011). *ED 6094 - Vous avez dit TMS ?*. Technical Report Edition INRS.
- Irmak, A., Bumin, G., & Irmak, R. (2012). The effects of exercise reminder software program on office workers' perceived pain level, work performance and quality of life. *Work: A Journal of Prevention, Assessment and Rehabilitation*, *41 Suppl 1*, 5692–5. doi:10.3233/WOR-2012-0922-5692.
- Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian Networks and Decision Graphs*. Information Science and Statistics. New York, NY: Springer New York. doi:10.1007/978-0-387-68282-2.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*, *22*, 5–55.
- Lim, Y.-K., Stolterman, E., & Tenenber, J. (2008). The anatomy of prototypes. *ACM Transactions on Computer-Human Interaction*, *15*, 1–27. doi:10.1145/1375761.1375762.
- Luo, S.-T., Su, C.-T., & Lee, W.-C. (2011). Constructing intelligent model for acceptability evaluation of a product. *Expert Systems with Applications*, *38*, 13702–13710. doi:10.1016/j.eswa.2011.04.162.
- Moitra, D. (1998). Managing change for software process improvement initiatives: a practical experience-based approach. *Software Process: Improvement and Practice*, *4*, 199–207. doi:10.1002/(SICI)1099-1670(199812)4:4<199::AID-SP107>3.0.CO;2-D.
- Nielsen, J. (1993). *Usability engineering*. San Diego, CA: Academic Press.
- Oniško, A., Druzdzal, M. J., & Wasyluk, H. (2001). Learning Bayesian network parameters from small data sets: application of Noisy-OR gates. *International Journal of Approximate Reasoning*, *27*, 165–182. doi:10.1016/S0888-613X(01)00039-1.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann.
- Roland Ortt, J., Langley, D. J., & Pals, N. (2007). Exploring the market for breakthrough technologies. *Technological Forecasting and Social Change*, *74*, 1788–1804. doi:10.1016/j.techfore.2007.05.009.
- Sauer, J., Seibel, K., & Rüttinger, B. (2010). The influence of user expertise and prototype fidelity in usability tests. *Applied ergonomics*, *41*, 130–40. doi:10.1016/j.apergo.2009.06.003.
- Sauer, J., & Sonderegger, A. (2009). The influence of prototype fidelity and aesthetics of design in usability tests: effects on user behaviour, subjective evaluation and emotion. *Applied ergonomics*, *40*, 670–7. doi:10.1016/j.apergo.2008.06.006.
- Silva, L. C. d. C. B., de Oliveira, A. B., Silva, D. C., Paschoarelli, L. C., & Coury, H. J. C. G. (2013). Evaluation of reusable cardboard box designs: Biomechanical and perceptual aspects. *International Journal of Industrial Ergonomics*, *43*, 154–160. doi:10.1016/j.ergon.2012.12.001.
- Sinclair, M. A. (2005). Subjective assessment. In E. N. Corlett, J. R. Wilson, & N. Corlett (Eds.), *Evaluation of Human Work* chapter 3. CRC Press. (3rd ed.).
- Song, B., Lee, C., & Park, Y. (2013). Assessing the risks of service failures based on ripple effects: A Bayesian network approach. *International Journal of Production Economics*, *141*, 493–504. doi:10.1016/j.ijpe.2011.12.010.
- Terrade, F., Pasquier, H., Reerink-Boulanger, J., Guingouain, G., & Somat, A. (2009). L'acceptabilité sociale : la prise en compte des déterminants sociaux dans l'analyse de l'acceptabilité des systèmes technologiques. *Le travail humain*, *72*, 383. doi:10.3917/th.724.0383.
- Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*, 186.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, *27*, 425–478.

Appendix A. Kolmogorov-smirnov test for the two cases

Table A.3: Descriptive parameters of normal distributions for the communicating pen case.

	FU	EC	GT	FCR	FA	MQ	CE	FB
Mean	4,111	1,722	1,667	4,167	1,722	3	4,222	3,222
Std. Deviation	0,7584	0,7519	0,767	0,7071	0,7519	1,414	0,8085	0,7321
Minimum	3	1	1	3	1	1	3	2
Maximum	5	3	3	5	3	5	5	4
First Quartile	3	1	1	3	1	2	3	2
Second Quartile	4	2	1,5	4	2	3	4	3
Third Quartile	4	2	2	4	2	4	5	3

Table A.4: Descriptive parameters of normal distributions for the active pause software case.

	PU1	PU2	PU3	PEOU1	PEOU2	FC1	FC2	SI1	SI2	ITU
Mean	4,929	4,964	4,786	5,929	5,893	5,393	5,321	4,429	4,393	4,607
Std. Deviation	1,412	1,374	1,969	1,331	1,197	1,37	1,335	1,476	1,499	2,233
Minimum	2	2	1	1	3	1	2	1	1	1
Maximum	7	7	7	7	7	7	7	7	7	7
First Quartile	4	4	3,5	5	4	4,5	3,5	2,5	2,5	1
Second Quartile	5	5	5,5	6	6	6	5	4	4	5
Third Quartile	6	5,5	6	7	7	6	6	5	5	6